Degree Project in Human-Computer Interaction
Second cycle, 30 credits

Avatar Control using Eye-Tracking in Virtual Reality
A user study on the sense of embodiment, user experience and workload using a Unity3D Prototype

AXEL GORIS
Avatar Control using Eye-Tracking in Virtual Reality

A user study on the sense of embodiment, user experience and workload using a Unity3D Prototype

AXEL GORIS

Degree Programme in Information and Communication Technology
Date: October 20, 2023

Supervisors: Christopher Peters, Fribourg Rebecca
Examiner: Tino Weinkauf
   School of Electrical Engineering and Computer Science
Host company: Ecole Centrale de Nantes
Swedish title: Avatar-kontroll med hjälp av ögonspårning i virtuell verklighet
Swedish subtitle: En användarstudie om känslan av förkroppsligande, användarupplevelse och arbetsbelastning med hjälp av en Unity3D-prototyp
Abstract

This master's thesis investigates methods for improving accessibility in virtual reality (VR), focusing on avatar control for individuals with limited mobility. VR offers an opportunity to help them with specific rehabilitation. Avatars, representing the user, serve as a bridge between the virtual and physical world, helping with the sense of embodiment and the more control, the better is the rehabilitation, for example. The objective is to identify effective avatar control approaches in VR, excluding traditional body movements, and assess their impact based on four criteria: sense of embodiment, user experience, task workload, and performance.

To address this, we designed a VR experiment in Unity3D with two tasks: one for interaction (selection and manipulation) and one for navigation. Each task had two independent variables. One common variable was the input modality (head-tracking or eye-tracking). For interaction, the second variable determined whether participants had to grasp an object to trigger the avatar's reaction or could directly control the avatar. In navigation, the second variable decided whether participants used steering or triggering mechanisms for navigating towards their destination. A user study using an HTC Vive Pro eye including 24 healthy participants, testing all four conditions for both tasks was conducted. Participants received tutorials before each condition. We recorded run times during each condition, with five runs for interaction and four for navigation. After each condition, participants provided feedback: the Virtual Embodiment Questionnaire assessed sense of embodiment, the NASA Task-Load Index evaluated task workload, and the System Usability Scale gauged user experience.

A repeated-measure two-way Anova showed that for interaction, eye-tracking improved speed performance but did not significantly affect the other metrics compared to head-tracking. In navigation, head-tracking consistently outperformed eye-tracking in all aspects. No other significant results were found. In conclusion, this thesis lays a foundation for enhancing VR accessibility by providing insights into input modalities and avatar control methods.

Keywords

Virtual Reality, Avatar Embodiment, Eye Tracking, User Experience, Accessibility
Sammanfattning


Nyckelord

Virtuell Verklighet, Avatar förkroppsligande, Ögonspårning, Användarupplevelse, Tillgänglighet
iv | Sammanfattning
Résumé

Ce mémoire de maîtrise étudie les méthodes permettant d’améliorer l’accessibilité de la réalité virtuelle (RV), en se concentrant sur le contrôle des avatars pour les personnes à mobilité réduite. La RV offre la possibilité de les aider dans le cadre d’une réadaptation spécifique. Les avatars, qui représentent l’utilisateur, servent de passerelle entre le monde virtuel et le monde physique, contribuant au sens de l’incarnation, et plus le contrôle est important, meilleure est la rééducation, par exemple. L’objectif est d’identifier des méthodes efficaces de contrôle des avatars dans la RV, à l’exclusion des mouvements corporels traditionnels, et d’évaluer leur impact sur la base de quatre critères : sens de l’incarnation, expérience de l’utilisateur, charge de travail et performance.

Pour répondre à cette question, nous avons conçu une expérience de RV dans Unity3D avec deux tâches : une pour l’interaction (sélection et manipulation) et une pour la navigation. Chaque tâche comportait deux variables indépendantes. Une variable commune était la modalité d’entrée (head-tracking ou eye-tracking). Pour l’interaction, la deuxième variable déterminait si les participants devaient saisir un objet pour déclencher la réaction de l’avatar ou s’ils pouvaient contrôler directement l’avatar. En ce qui concerne la navigation, la deuxième variable déterminait si les participants utilisaient des mécanismes de direction ou de déclenchement pour naviguer vers leur destination. Une étude d’utilisateur utilisant un HTC Vive Pro Eye comprenant 24 participants en bonne santé, testant les quatre conditions pour les deux tâches a été menée. Les participants ont reçu des tutoriels avant chaque condition. Nous avons enregistré les temps d’exécution pour chaque condition, avec cinq exécutions pour l’interaction et quatre pour la navigation. Après chaque condition, les participants ont donné leur avis : le Virtual Embodiment Questionnaire a évalué le sentiment d’incarnation, le NASA Task-Load Index a évalué la charge de travail, et le System Usability Scale a mesuré l’expérience de l’utilisateur.

Un repeated-measure two-way Anova a montré que pour l’interaction, l’eye-tracking a amélioré les performances en termes de vitesse mais n’a pas eu d’effet significatif sur les autres mesures par rapport au head-tracking. En ce qui concerne la navigation, le head-tracking a été plus performant que l’eye-tracking dans tous les domaines. Aucun autre résultat significatif n’a été trouvé. En conclusion, cette thèse pose les bases de l’amélioration de l’accessibilité à la RV en fournissant des informations sur les modalités d’entrée et les méthodes de contrôle des avatars.
Mots-clés

Réalité Virtuelle, Incarnation d’Avatar, Suivi Oculaire, Expérience Utilisateur, Accessibilité
Acknowledgments

My sincere thanks to Rebecca Fribourg, Jean-Marie Normand and Geoffrey Gorisse for having organized this research as well as helped me define the project and find a coherent research question. They have been a great support as well as very nice supervisors who helped me challenge my assumptions about what I thought.

I would also like to thank Christopher Peter for accepting to supervise my thesis and help me focus on the important parts.

I would like to thank the "Ambiances-Architectures-Urbanités" lab for providing me with all the required material.

My deepest thanks to Théo Combe for being present almost all along my master thesis and being a person I could rely on, whether for ranting about Unity or for talking about what I wanted to do. Thanks a lot to Martin Guy and Nicolas Fourrier as well for all our discussions and your good advices!

Rémi Boutroux, thank you for testing my application an absurd amount of time, and I’m sorry for making you try very unpleasant navigation methods. Thanks a lot for having been here!

I could go on forever, but in no particular order, I would like to thank Jean-Yves Martin for sharing lots of surprising stories with me, Hafsa and Sarah for teaching me some Arabic words, Nöé for no particular reason and Flora and Louis for having me as a guest once a week. Thanks everyone! :)

Nantes, France, October 2023
Axel Goris
viii | Acknowledgments
# Contents

## 1 Introduction
1.1 Background ............................................ 1
1.2 Problem ................................................. 2
  1.2.1 Original problem and definition ................... 3
  1.2.2 Scientific and engineering issues ................. 5
1.3 Purpose .................................................. 6
1.4 Benefits, Ethics, and Sustainability .................. 6
1.5 Research Methodology ................................ 7
1.6 Delimitations .......................................... 8
1.7 Structure of the thesis ................................ 9

## 2 Background
2.1 Definitions .............................................. 11
  2.1.1 Virtual Reality and Virtual Environment .......... 11
  2.1.2 Avatar Representation in Virtual Environments ... 13
  2.1.3 Embodiment and Sense of Embodiment ............... 13
  2.1.4 Accessibility ......................................... 16
  2.1.5 Categorization of 3D Interaction Techniques ...... 17
  2.1.6 Eye-Tracking ......................................... 21
  2.1.7 Evaluation ............................................ 22
2.2 Related Work ........................................... 23
  2.2.1 Manipulation of the Sense of Agency ............... 23
  2.2.2 Hands-free interaction ............................... 25
  2.2.3 Eye-Tracking for Selection .......................... 26
  2.2.4 Eye-Tracking for Manipulation ...................... 28
  2.2.5 Eye-Tracking for Navigation ........................ 29
2.3 Summary .................................................. 30
# Contents

3 Methods .......................................................... 33
  3.1 Research Process ........................................... 33
  3.2 Experimental Setup ....................................... 34
    3.2.1 Pick and Place .................................... 36
    3.2.2 Navigation ......................................... 39
  3.3 Data Collection ............................................. 42
    3.3.1 Sampling ........................................... 42
  3.4 Experimental design/Planned Measurements ............... 43
    3.4.1 Hardware/Software used ............................ 43
  3.5 Assessing reliability and validity of the data collected .. 44
    3.5.1 Reliability Assessment ............................ 44
    3.5.2 Validity Assessment ............................... 44
    3.5.3 Validity of method ................................ 45
  3.6 Data Analysis ............................................... 47
  3.7 Conclusion .................................................. 49

4 Results and Analysis ........................................... 50
  4.1 Results ..................................................... 50
    4.1.1 Selection & Manipulation ........................... 50
    4.1.2 Navigation .......................................... 53
  4.2 Impact of Virtual Reality and Video game Habit ........... 57
  4.3 Conclusion .................................................. 59

5 Discussion ...................................................... 61
  5.1 Research Question 1: Selection and Manipulation ........... 61
    5.1.1 Performance Improvement with Eye-Tracking ........ 61
    5.1.2 Selection and Manipulation Performance with Avatar .. 62
    5.1.3 Task Duration ..................................... 62
    5.1.4 Level of Experience ................................ 62
    5.1.5 Comparison with Previous Studies ................. 62
    5.1.6 Conclusion ......................................... 63
  5.2 Research Question 2: Navigation ............................. 63
    5.2.1 Navigation Task Differences ....................... 63
    5.2.2 Navigation Methods ................................ 63
    5.2.3 Rotation in Navigation .............................. 64
    5.2.4 Level of experience ................................ 64
    5.2.5 Comparison with Previous Studies ................. 64
    5.2.6 Conclusion ......................................... 64
  5.3 Overall Implications and Broader Context .................... 64
5.3.1 Sense of Embodiment in VR ........................................ 65
5.3.2 User Experience and Task Performance ..................... 65
5.3.3 Broader Implications and Societal Impact .................... 65
5.4 Conclusion ................................................................ 66
5.5 Limitations ................................................................. 66
  5.5.1 Avatar Realism and Implications ............................... 66
  5.5.2 Healthy Population ................................................ 67
  5.5.3 Participant Demographics ....................................... 68
  5.5.4 No baseline .......................................................... 68
  5.5.5 Eye-Tracking Challenges ........................................ 68
  5.5.6 Level of Experience .............................................. 68
  5.5.7 Conclusion ........................................................... 69
6  Conclusions and Future work ............................................ 71
  6.1 Conclusions ............................................................... 71
  6.2 Future work ............................................................. 72
    6.2.1 Baseline Comparison ........................................... 72
    6.2.2 Adaptation to Existing Software ............................ 72
  6.3 Reflections ............................................................... 72
References ....................................................................... 73
A  Questionnaires ............................................................... 87
  A.1 Virtual Embodiment Questionnaire .............................. 88
  A.2 System Usability Scale ............................................. 89
  A.3 Nasa Task-Load-Index .............................................. 91
  A.4 Demographics Questionnaire ................................... 92
B  Data Graphs and Tables .................................................... 93
  B.1 Performance ............................................................ 93
  B.2 Sense of Embodiment .............................................. 99
  B.3 System Usability .................................................... 104
  B.4 Nasa Task-Load-Index ............................................ 107
  B.5 Virtual Reality and Video game Experience Impact ....... 109
List of Figures

2.1 Perception-Action loop involving avatars. Real users embody a virtual avatar that represents them in the virtual environment. They interact with the Virtual Environment (VE) through that avatar and receive feedback that in return improve their experience and increase their sense of embodiment. Considering this is only a prototype, no other user will be involved. Heavily inspired by the work from Rebecca Fribourg. 12

3.2 The distinction between input modality and the methods used for the selection & manipulation task and the navigation task. 35

3.3 The two avatars used for both tasks. Participants were attributed one at the beginning of the experiment and would keep the same during the whole experiment. 36

3.4 The virtual environment with the avatar for the interaction task where the user was building a burger from the user point of view and a third point of view. A complete video showcasing the four conditions including auditory feedback as well as the waiting room between conditions can be found at https://youtu.be/XXdqE9XEbEl 37

3.5 The virtual environment and the avatar for the navigation task where the user was going from one cube to another. The two cubes with outlines are the two cubes for rotating to the side. The small yellow sphere on the ground indicates where the user is pointing. A complete video showcasing the four conditions including auditory feedback can be found at https://youtu.be/L4ekN0GMF9c 40

3.1 The timeline that was followed in this master thesis, presenting what was implemented and what was done, for the development, for the user testing and for the research process in general. 48
4.1 The time it took to complete a run, on average for each participant, depending on whether it was the first, second, third, fourth, or fifth run of the condition for the interaction task. The first run was systematically excluded considering the time it took to complete was way higher than the rest of the run, probably due to an unclear starting signal. For the signification of “*” see 4.1.

4.2 The time it took on average to complete a run for the interaction task (taking an ingredient, putting it in the plate) depending on the two independent variables, Degree of Control (see 3.2.1.1) and Input Modality (see 3.2.1.2). For the signification of “*” see 4.1.

4.3 The time it took on average to complete a run in the navigation task (going from one glowing cube to another glowing cube) depending on the two independent variables, Navigation Method (see 3.2.2.1) and Input Modality (see 3.2.1.2). For the signification of “*” see 4.1.

4.4 The sense of agency score in the navigation task, referred to as the control score here, one of the subcomponent of the Sense of Embodiment (SoE), measured using the Virtual Embodiment Questionnaire (VEQ) and displayed depending on the Input Modality. For the signification of “*” see 4.1.

4.5 System Usability Scale (System Usability Scale (SUS)) score depending on the Input Modality for the navigation task. For the signification of “*” see 4.1.

4.6 Nasa Task-Load-Index (Nasa Task Load Index (NASA-TLX)) Score depending on the Input Modality for the navigation task. For the signification of “*” see 4.1.

4.7 Time taken for a run in the selection & manipulation task, depending on VR and Video game Experience. The Virtual Reality (VR) experience seems to be playing a role in the results, considering that people who never tried VR have worse results than others. For the signification of “*” see 4.1.

4.8 Time taken for a run in the navigation task, depending on VR and Video game Experience. For the signification of “*” see 4.1.

A.1 Virtual Embodiment Questionnaire

A.2 System Usability Scale Questionnaire
A.3 Nasa Task-Load-Index Questionnaire ........................................... 91
A.4 Demographics Questionnaire used before the experiment ....... 92
B.1 Interaction Run Duration Depending on the coupled conditions 94
B.2 Interaction Total Duration Depending on the coupled conditions 95
B.3 Navigation Run Duration depending on the order of the run ... 96
B.4 Navigation Run Duration depending on the rotation direction . 97
B.5 Navigation Run Duration depending on the Coupled Conditions 98
B.6 Navigation Total Duration depending on the Coupled Conditions 99
B.7 Acceptance depending on the Degree of Control and Input
    Modality .................................................................................. 100
B.8 Control depending on the Degree of Control and Input Modality 101
B.9 Change depending on the Degree of Control and Input Modality 101
B.10 Acceptance depending on the Navigation Method and Input
    Modality .................................................................................. 102
B.11 Control depending on the Navigation Method and Input Modality 103
B.12 Change depending on the Navigation Method and Input Modality 103
B.13 System Usability depending on the Degree of Control and
    Input Modality ......................................................................... 104
B.14 System Usability depending on the Coupled Conditions ....... 105
B.15 System Usability depending on the Navigation Method and
    Input Modality ......................................................................... 106
B.16 System Usability depending on the Coupled Conditions ....... 107
B.17 Task Load Index depending on the Degree of Control and
    Input Modality ......................................................................... 108
B.18 Task Load Index depending on the Navigation Method and
    Input Modality ......................................................................... 109
xvi | List of Figures
List of Tables

4.1 In every graph, a "*" between two conditions represents the p-value. If no star is present, then the results are not significant. 50

4.2 p-values for interaction performance (time taken for a run) based on the target and method used. 52

4.3 p-values for navigation performance based on the target and method used. 54

B.1 Tukey - Interaction - VR Table 109
B.2 Tukey - Interaction - Video game Table 110
B.3 Tukey - Navigation - VR Table 110
B.4 Tukey - Navigation - Video game Table 110
xviii | List of Tables
# List of acronyms and abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCI</td>
<td>Brain-Computer Interface</td>
</tr>
<tr>
<td>EEG</td>
<td>Electroencephalogram</td>
</tr>
<tr>
<td>EMG</td>
<td>Electromyography</td>
</tr>
<tr>
<td>HCI</td>
<td>Human Computed Interaction</td>
</tr>
<tr>
<td>HMD</td>
<td>Head-Mounted Display</td>
</tr>
<tr>
<td>IK</td>
<td>Inverse Kinematics</td>
</tr>
<tr>
<td>NASA-TLX</td>
<td>Nasa Task Load Index</td>
</tr>
<tr>
<td>SoA</td>
<td>Sense of Agency</td>
</tr>
<tr>
<td>SoBO</td>
<td>Sense of Body-Ownership</td>
</tr>
<tr>
<td>SoE</td>
<td>Sense of Embodiment</td>
</tr>
<tr>
<td>SoSL</td>
<td>Sense of Self-Location</td>
</tr>
<tr>
<td>SUS</td>
<td>System Usability Scale</td>
</tr>
<tr>
<td>UI</td>
<td>User Interface</td>
</tr>
<tr>
<td>VE</td>
<td>Virtual Environment</td>
</tr>
<tr>
<td>VEQ</td>
<td>Virtual Embodiment Questionnaire</td>
</tr>
<tr>
<td>VR</td>
<td>Virtual Reality</td>
</tr>
</tbody>
</table>
List of acronyms and abbreviations
Chapter 1

Introduction

In recent years, video games have emerged as a highly influential medium, with an estimated 2.7 billion individuals engaging in gaming activities by 2023 [1]. With the game industry’s significant investment in Virtual Reality (VR) technology and its widespread adoption among professionals, it has become imperative to address the accessibility challenges faced by individuals with disabilities in VR environments [2, 3, 4]. While current VR experiences predominantly rely on hand-based interactions, it is crucial to consider the needs of individuals who lack control over their hands or have limited hand dexterity.

This thesis aims to explore the feasibility and effectiveness of utilizing eye-tracking technology for controlling VR avatars. By leveraging eye-tracking as an alternative input modality, we aim to empower individuals with limited hand control to interact seamlessly within VR environments. By examining the potential of eye-tracking as a control method, we can contribute to the development of more inclusive and accessible VR experiences for a diverse range of users.

1.1 Background

VR has become a popular medium for immersive experiences, typically involving the use of Head-Mounted Display (HMD) and motion controllers. In these virtual environments, users are often represented by virtual bodies, known as avatars, which can be controlled through body movements or motion controller inputs. However, individuals with disabilities, such as quadriplegia, face challenges in accessing and interacting with VR content, as they may not have the physical capability to use body tracking or motion controllers
effectively. This limitation prevents them from fully experiencing the potential benefits of VR immersion and embodiment.

Moreover, VR technology has found applications in the medical field, particularly in rehabilitation settings, where patients can engage in virtual training and therapy sessions to enhance their recovery process [5]. Brain-Computer Interface (BCI) have been employed as a means of control for individuals with disabilities, enabling them to interact with VR environments. However, BCI come with certain drawbacks, including high costs, extensive training requirements, and varying degrees of success in achieving effective control [6].

In this context, eye-tracking technology holds promise as an alternative and potentially more accessible method for controlling avatars in VR for individuals with disabilities. By providing a user-friendly and intuitive means of control, eye-tracking can offer enhanced embodiment experiences, particularly in terms of the Sense of Agency (SoA) - the feeling of being in control of one’s avatar’s movements, which is a crucial aspect of the Sense of Embodiment (SoE) [7]. As highlighted by [8], exploring the use of eye-tracking technology can empower individuals with severe motor impairments, such as quadriplegics or individuals who are locked in, to navigate their avatars in VR.

1.2 Problem

In the United States alone, it is estimated that 291,000 individuals are living with spinal cord injuries, with approximately 60% of them classified as quadriplegic [9]. Extrapolating this rate of incidence to a global scale suggests that there are approximately 4,800,000 quadriplegic individuals worldwide due to spinal cord injuries. These individuals face significant challenges in accessing and enjoying the vast array of VR experiences that are currently available. Furthermore, it is important to note that this number only represents a portion of the overall population with disabilities, as it does not account for other sources of impairment.

Given the potential benefits of avatars in rehabilitation contexts, it is crucial to address the lack of accessibility and provide appropriate alternatives for individuals with disabilities. By enabling access to immersive VR experiences, individuals with limited physical capabilities can benefit from the engaging and interactive nature of virtual environments, enhancing their rehabilitation process and overall well-being.
1.2.1 Original problem and definition

In this study, we aim to explore the use of eye-tracking as a potential solution for avatar controls in VR. Eye-tracking technology holds promise as an accessible input method for individuals with quadriplegia, offering a means of interaction that does not rely on physical movement. However, it is important to acknowledge that certain eye conditions, such as strabismus or squint, may pose limitations for some individuals.

1.2.1.1 Research Question

RQ: How can different eye-tracking based control methods of an avatar impact users’ sense of embodiment and more globally the user experience in VR?

This research question could be divided into multiple questions. Considering the broad scope of the research question, we think it is better to subdivide it into smaller research questions.

Research Question 1: How will the different methods of eye-tracking selection and manipulation impact the SoE over the avatar, and how will this in turn affect the usability of our eye-tracking control methods?

The eye tracking technology used in this study provides a ray into the scene, enabling users to select objects by directing their gaze towards them.

While various selection methods (such as dwell, wink, voice, dual gaze) have been previously examined \cite{10, 11, 12, 13}, they do not fundamentally alter the way users interact with an avatar. In this study, we aim to compare two distinct methods of avatar interaction:

- Condition 1: In this condition, the selection of an object triggers the avatar’s movement. For example, if the user selects a cube, the avatar will grab the cube. However, the user does not have direct control over the avatar’s movement to reach the selected object.

- Condition 2: In this condition, the movement of the avatar triggers the selection of an object. The user first selects the avatar’s hand and moves it to the cube to select it, providing greater control over the avatar’s movement.

In comparing these two conditions, it is important to consider the trade-off between control and performance. While one method may offer more control over the avatar’s movement, it could potentially be more tiring, resulting in
reduced usability, desirability, and performance. To assess performance, we will measure the time taken to complete the selection task in each condition, providing a rough estimate of the performance of each control method. To assess usability, participants will be asked to complete the System Usability Scale (SUS) questionnaire [14]. Additionally, participants will be asked to complete the NASA Task Load Index (NASA-TLX) questionnaire [15] to provide subjective ratings of mental workload.

Furthermore, to evaluate the SoE experienced by participants, we will employ the Virtual Embodiment Questionnaire (VEQ) [16]. The VEQ measures the extent to which participants feel embodied in the virtual environment, encompassing various aspects such as body-ownership, agency, and self-location.

By addressing these research questions and utilizing the VEQ, NASA-TLX, and SUS measures, we aim to gain insights into the impact of different eye-tracking-based control methods on users’ sense of embodiment and overall user experience in virtual reality. The results can be found in 5.1.

To ensure a comprehensive virtual reality (VR) experience, navigation plays a crucial role as it is present in almost every VR scenario [17].

**Research Question 2: How do different methods of navigation influence users’ sense of embodiment and usability?**

Previous studies have explored the effectiveness of teleportation using head-tracking with wink or voice commands [12, 18]. However, to avoid the Midas Touch problem [19], we will not employ dwell or dual gaze selection methods on the floor, as they can lead to poor user experience and interface complications. In line with Research Question 1, we will compare two distinct methods of avatar navigation:

- **Condition 3: Trigger Navigation** - In this condition, the user selects a point on the floor, and upon a specific command (e.g., long blink), the avatar initiates navigation towards the selected point.

- **Condition 4: Continuous Navigation** - In this condition, the user can switch between navigation and non-navigation modes using a command (e.g., long blink). When in navigation mode, the avatar moves towards the user’s gaze target.

Similar to Research Question 1, we will employ the same questionnaire to assess the SoE experienced by participants after they complete a navigation task. The SUS and NASA-TLX will also be used. The time taken to complete
the navigation task will serve as a rough estimate of the performance of each navigation method. The results can be found in 5.2.

Given that many eye-tracking control methods can be adapted for head orientation, resulting in reduced eye fatigue, there is a possibility that head-tracking could serve as an effective alternative to eye-tracking for individuals with quadriplegia. Considering that individuals with quadriplegia often retain control over their head movements, it becomes important to assess the feasibility and effectiveness of head-tracking in comparison to eye-tracking. Furthermore, exploring the potential inclusion of individuals with eye diseases in the context of head-tracking provides valuable insights into the broader applicability of this control method.

**Research Question 3:** Can head-tracking yield superior results compared to eye-tracking for individuals who retain control over their head movements?

Previous research has demonstrated that head-tracking can exhibit both superior and inferior performance compared to eye-tracking, depending on the specific use case [20, 21]. In our study, we aim to provide the optimal interaction experience for users, considering their individual abilities. Consequently, we will compare each selection, manipulation, and navigation method utilizing eye-tracking with their counterparts using head-tracking. The results can be found in 5.1 and 5.2.

### 1.2.2 Scientific and engineering issues

The utilization of eye-tracking as a sole input modality for controlling a comprehensive range of interactions within a VR environment presents several scientific and engineering challenges. The available modalities are the following:

- **2D plane:** Eye-tracking primarily operates within a 2D plane, where the user’s gaze is directed towards the VR scene. While this poses minimal issues for selection tasks, accurately mapping eye movements for manipulation tasks, which involve interactions in three-dimensional space, becomes a complex challenge. Determining the depth at which the user intends to interact within the VR environment solely based on eye-tracking data can be problematic.

- **Triggers:** Eye states (open, closed, half-closed) and blink durations, can be used to map multiple buttons. However, this type of interaction can be mentally and physically exhausting for users, as well as potentially...
leading to unintended actions. This raises concerns regarding the usability and efficiency of eye-tracking as an interaction modality within a VR context, as highlighted in previous research [22]. The Midas Touch problem, where unintentional actions occur due to accidental triggers, is an additional challenge that needs to be addressed [19].

Another significant issue pertains to navigation within the VR environment using an avatar representation. Maintaining a sense of embodiment and presence is crucial for a positive user experience and sense of ownership over the avatar. Abrupt avatar disappearance or teleportation can disrupt this embodiment. To mitigate this issue, navigation techniques will be implemented in a manner that gradually transitions the user’s point of view to the destination. However, this gradual movement can induce discomfort or simulation sickness in some users, potentially compromising the validity of the results. To address this, several strategies will be implemented, including reducing the field of view during movement, employing slow acceleration and deceleration when starting or stopping motion, and maintaining a moderate speed during movement to minimize disorientation.

By recognizing and addressing these scientific and engineering challenges associated with eye-tracking-based control methods and avatar navigation, we can enhance the usability, comfort, and overall user experience in VR environments for individuals with quadriplegia.

1.3 Purpose

The primary objective of this research is to identify a method of controlling an avatar in VR without relying on body movements, and to evaluate its performance and user experience, specifically in terms of the SoE. In order to achieve this goal, a comparative analysis of multiple control methods will be conducted.

1.4 Benefits, Ethics, and Sustainability

The project aligns with Goal 10 of the United Nations’ 17 Sustainable Development Goals, which focuses on reducing inequalities. Specifically, it contributes to achieving Goal 10.2 [23], which aims to promote the social, economic, and political inclusion of all individuals by addressing the barrier of entry in VR for people with disabilities. By developing accessible control methods for avatars in VR that do not rely on full-body control, the project
aims to remove inequalities and provide equal opportunities for individuals with limited mobility, such as those with quadriplegia.

In terms of ethical considerations, the project takes into account the use of eye-tracking technology, which enables the collection of physiological measures in response to different situations. To ensure ethical conduct, the research will prioritize user privacy and data protection. The project will not store any sensitive or identifiable physiological data without the explicit consent of the participants. Transparency will be maintained by informing users about the data being collected and stored, and measures will be in place to protect user confidentiality. Moreover, before proceeding to the user experiment, Centrale Nantes’ research committee did approve the experiment.

Regarding sustainability, the project strives to minimize energy consumption in the design of the control methods. Although VR technology itself may have some environmental impact, the project’s focus on enabling rehabilitation in VR, rather than relying on larger physical simulators or expensive equipment, is a step towards sustainability. By leveraging VR as a rehabilitation tool, the project aims to reduce the need for extensive physical infrastructure, ultimately contributing to a more sustainable approach to healthcare and accessibility.

Overall, the project acknowledges the importance of addressing social inequalities, ensuring ethical practices in data collection, and striving for sustainability in the context of VR avatar control for individuals with disabilities. By actively working towards these considerations, the project aims to create a positive impact in line with the United Nations’ Sustainable Development Goals.

### 1.5 Research Methodology

Based on the research methods and methodologies portal by Hakansson [24], our study adopts a quantitative research approach with measurable hypotheses. We align with the philosophical assumption of Positivism, particularly because our research involves user testing within the field of Human Computed Interaction (HCI). By ensuring that the measures are independent of the observers and instruments, we aim to maintain objectivity and empirical rigor.

To investigate the impact of different control commands on user interaction, we will employ an experimental research design. Through the manipulation of control commands and measurement of outcomes, we can verify or falsify hypotheses in a deductive manner. The experiment will take place in a VR environment, where we have control over various factors that
could potentially influence the experimental results. This controlled approach allows us to systematically analyze the effects of different control methods on user experience.

In summary, our research approach integrates quantitative methods, adheres to positivist philosophical assumptions, and utilizes an experimental research design within the domain of HCI. Through this framework, we seek to examine how variations in control commands impact user interaction, employing VR as the experimental setting while controlling relevant variables for a robust investigation.

1.6 Delimitations

This project focuses on specific aspects of VR interaction and does not encompass certain areas. The following aspects will not be directly addressed:

- Selection methods: The project does not involve comparing multiple input methods such as Wink [12], Dwell, DualGaze [25] and Outline Pursuits [26]. These methods have been explored in previous studies, but are not the primary focus of this research.

- Voice selection: Despite its potential, voice selection will not be examined due to its limited applicability in social rehabilitation situations where voice is already utilized for communication.

- Complex manipulation: The project will not delve into complicated manipulation tasks like rotation or scaling [27]. Instead, the aim is to explore possibilities for disabled individuals in VR using avatars, focusing on simpler interactions.

- Animation vs. inverse kinematics [28]: The study will employ inverse kinematics for handling hand movements, but will not compare different animation methods.

- Sense of Body-Ownership and its relation to interaction: While the evaluation of the Sense of Body-Ownership will be conducted, the project will not actively attempt to modify or influence it.

- Real-life situations vs. clinic tests: The study acknowledges that the repetitive nature of clinic tests may limit the direct applicability of the findings to real-world scenarios.
These exclusions allow for a focused investigation into specific aspects of avatar-based VR interaction for disabled individuals, providing insights and understanding within the scope of the project’s objectives.

1.7 Structure of the thesis

This thesis first performs a short state-of-the-art dealing with avatars, the SoE, different class of interaction and input modalities, more specifically about eye-tracking. Then a new technique and an existing one are selected for selection and two existing ones are selected for navigation and their processes are fully explained. Part 3 contains details about the two experiment that are conducted. Part 4 presents and discuss obtained results after conducting user testing. Finally, a discussion is proposed in part 5 and part 6 presents a conclusion and future work are discussed.
Chapter 2

Background

This chapter provides basic background information about HCI and the SoE.

2.1 Definitions

This section of the master thesis presents the essential definitions and delineations of the various areas of study. It serves to establish a conceptual framework and provide clarity on the specific aspects that will be explored within the research project.

2.1.1 Virtual Reality and Virtual Environment

VR technology enables users to experience immersion within Virtual Environment (VE) while actively interacting with the digital content, which can include virtual objects or other users. VE are three-dimensional representations that are dynamically computed and displayed in real-time as the user navigates and explores the virtual space. The primary aim of VR is to create a sense of presence, where users feel as though they are physically present within the virtual world.

To achieve this immersive experience, VR often utilizes HMD. These devices are worn on the user’s head and provide visual and auditory feedback, effectively blocking out the physical environment and replacing it with the virtual content. HMD typically consist of high-resolution displays positioned close to the user’s eyes, offering a wide field of view and enhancing the sense of immersion. They may also incorporate integrated headphones or speakers to provide spatial audio, further enhancing the sense of presence and realism.
While visual and auditory feedback are the primary modes of interaction in VR, haptic feedback can also be employed, although it is less commonly utilized. Haptic feedback refers to the tactile and force-based feedback that allows users to sense and interact with virtual objects or receive physical feedback in response to their actions. Examples of haptic feedback in VR include vibration or force feedback in handheld controllers or specialized gloves that provide a sense of touch and manipulation within the virtual environment. However, due to technical limitations and cost considerations, haptic feedback is not as prevalent as visual and auditory feedback in most VR experiences.

In summary, VR technology offers users the opportunity to immerse themselves in virtual environments and interact with digital content, as shown in figure 2.1.

![Figure 2.1: Perception-Action loop involving avatars. Real users embody a virtual avatar that represents them in the virtual environment. They interact with the VE through that avatar and receive feedback that in return improve their experience and increase their sense of embodiment. Considering this is only a prototype, no other user will be involved. Heavily inspired by the work from Rebecca Fribourg.](image)

VR is important in this master thesis, considering it is the primary user interface.
2.1.2 Avatar Representation in Virtual Environments

In many applications, users are represented in the VE by a virtual body, commonly referred to as an avatar, which is animated in real time based on their own movements. This representation allows users to embody a virtual entity and interact with the VE using body actions. Various techniques can be employed to achieve real-time avatar animation, such as motion capture suits or specific sensors (e.g., head, hands, and feet sensors) combined with inverse kinematics algorithms [28]. These approaches enable the estimation and mapping of body movements onto the virtual avatar, creating a sense of embodiment and agency.

It is important to note that the general definition of an “avatar” does not necessarily involve a body. For instance, Sherman and Craig [29] defined an avatar in 2002 as follows:

“A virtual object used to represent a participant or a physical object in a virtual world; the (typically visual) representation may take any form.”

This broader definition acknowledges that avatars can encompass various visual representations beyond humanoid bodies, allowing for flexibility and creativity in their design and utilization within virtual worlds. The avatar representation and the way in which it is animated are important to this work, considering it can drastically improve or degrade the feeling of the sense of embodiment.

2.1.3 Embodiment and Sense of Embodiment

In this paper, we differentiate between embodiment and the Sense of Embodiment (SoE). Embodiment can be defined as the processing of some properties of an entity (E) in the same way as the properties of one’s own body, as stated by De Vignemont [30]. It encompasses the cognitive and neural mechanisms involved in the integration of external entities with one’s own body representation.

On the other hand, the SoE refers to the subjective perception of embodiment, specifically how the process of embodiment is experienced by the user. According to Kilteni [7], the SoE encompasses all the sensations associated with being, having, and controlling a virtual body. It can be further divided into three components: the Sense of Body-Ownership, the Sense of Agency, and the Sense of Self-Location. These components capture different aspects of the subjective experience of embodiment in virtual reality.

Alternative definitions of the SoE exist, such as the one proposed by De Vignemont [30] or by Roth [31]. However, for the purpose of this thesis, we
will adopt the definition provided in [7].

### 2.1.3.1 The Sense of Body-Ownership

The Sense of Body-Ownership (SoBO) can be thought of as the subjective experience of perceiving a virtual body as “This is my body”. It can refer to one part or to a full body. By establishing sensory connections between one’s biological body and the stimulus perceived on the virtual body in virtual reality, such as visuotactile or visuoproprioceptive correlations, one might raise the Sense of Body-Ownership (SoBO). The well-known rubber-hand illusion is based on this idea [32]. Basically, the more a person’s body and virtual body look alike, the more SoBO is perceived.

In VR, various techniques have been employed to enhance the SoBO, including realistic body avatars, haptic feedback, and multisensory integration. For example, by providing a virtual body that closely matches the user’s physical appearance and movements, the sense of body-ownership can be significantly strengthened [7].

Understanding the factors that contribute to the SoBO in VR is crucial for creating more immersive and engaging virtual experiences. By leveraging the principles of visuotactile and visuoproprioceptive correlations, researchers, and developers can design virtual environments that foster a stronger sense of body-ownership and increase the user’s embodiment in the virtual body.

### 2.1.3.2 The Sense of Agency

The Sense of Agency (SoA) can be understood as the subjective experience of being the causal agent or generator of an action [33]. In the context of this study, the SoA is of particular importance as we explore different interaction methods in VR.

The SoA can be further divided into two components [33]: (1) the feeling of agency and (2) the judgment of agency. The judgment of agency is computed by comparing the predicted outcome of an action with the actual outcome, and it occurs after the feedback has been perceived and processed. On the other hand, the feeling of agency arises from the perception of feedback at the early stages of the action.

To perceive a judgment of agency, three principles, as outlined by Wegner [34], must be respected:

- **The priority principle**: The intention to perform an action must precede the action itself, which in turn must directly precede the outcome.
• **The consistency principle:** The actual outcome should be similar to the predicted outcome.

• **The exclusivity principle:** The action must be fully attributed to oneself and not influenced by external factors.

These principles play a crucial role in shaping the sense of agency and determining whether individuals perceive themselves as the causal agents of their actions in the virtual environment.

By investigating and understanding the factors that influence the SoA, we can design and implement interaction methods in VR that enhance the user’s sense of agency, ultimately leading to more immersive and engaging virtual experiences. This will be the primary mean of action towards the SoE in this thesis, considering that we are modifying the way in which users interact and control the avatar.

### 2.1.3.3 The sense of Self-Location

The Sense of Self-Location (SOSL) can be conceptualized as the subjective experience of being inside a body. It refers to a specific volume in space where an individual perceives themselves to be located. Unlike the simple spatial experience of being inside the environment, Sense of Self-Location (SoSL) specifically relates to the spatial experience of being within a body. This perception is heavily influenced by the egocentric visuospatial perspective.

In the context of VR, SoSL can be experienced from a first-person perspective by attending to the position of the artificial body’s eyes. This perspective allows users to feel a sense of self-location within the virtual environment, as if they were truly inside the represented body. Conversely, adopting a third-person viewpoint disrupts the normal circumstances in which participants typically feel self-location in relation to their actual body [35, 36].

By understanding the factors that contribute to SoSL, researchers and developers can design VR experiences that promote a stronger sense of presence and embodiment, ultimately enhancing the user’s overall immersive experience. This should not change in any conditions in our work.

### 2.1.3.4 Evaluation of the Sense of Embodiment

Evaluating the SoE in VR can be a challenging task due to its subjective nature. Subjective measures, such as questionnaires, have been commonly used to assess the SoE [37, 16]. Questionnaires aim to standardize the evaluation of the SoE; however, it is important to consider that the interpretation of
questionnaire items may vary among users. Additionally, administering questionnaires after the experiment may overlook the potential impact of time or specific events on participants’ responses [38].

Alternatively, objective measures, such as neurophysiological markers, can be utilized [39]. However, these measures can be complex to set up and often require expensive equipment. One common objective measure involves introducing a virtual threat to the avatar and observing the user’s reaction. It has been shown that if the user reacts to a virtual threat, it corresponds to a strong sense of embodiment [40]. Nevertheless, it is worth noting that introducing a threat may influence the subjective measure of the SOE [41]. Furthermore, in experiments with multiple repetitions, the efficacy of threat introduction may decrease over time, leading to reduced participant response for within-subject comparisons.

In conclusion, evaluating the SoE in VR involves a balance between subjective and objective measures. Subjective measures, such as questionnaires, provide insights into users’ perceptions, but careful consideration must be given to questionnaire design and administration timing. Objective measures, although more challenging to implement, offer a potentially more reliable assessment of the SoE. Both approaches contribute to a comprehensive evaluation of the SoE and help advance our understanding of the user experience in VR.

2.1.4 Accessibility

Users with limited mobility or physical disabilities pose unique challenges for VR applications. However, advancements in inclusive design and medical solutions have made it possible to cater to these individuals’ needs and provide them with meaningful VR experiences.

In recent years, there has been a growing emphasis on inclusivity in VR applications, ensuring accessibility for people with conditions such as paraplegia or tetraplegia who are unable to move their own bodies. VR has shown promise in providing medical solutions for stroke patients with limited limb control, aiding in their limb rehabilitation process [42].

While some VR applications provide visual feedback without user control, such as presenting animations of a virtual body walking from a first-person perspective, other applications strive to incorporate user control despite their lack of mobility. Even a minimal amount of control over a virtual body can significantly enhance the sense of embodiment and positively impact rehabilitation and overall VR experience [7].
Researchers have explored various approaches to enable user control over virtual bodies, including the use of Electroencephalogram (EEG)-based brain-computer interfaces (BCI). These systems allow users to exert control over their virtual bodies through their brain activity, enhancing their sense of agency and ownership. Studies have demonstrated the positive effects of such systems in inducing feelings of control and improving the overall VR experience [43].

However, utilizing EEG-based BCI in VR applications presents several technical challenges and requires extensive training and installation procedures [6]. These challenges include ensuring accurate and real-time decoding of brain signals, addressing individual differences in brain activity patterns, and establishing reliable and user-friendly interfaces.

Despite these challenges, the potential benefits of incorporating user control for individuals with limited mobility in VR are significant. By continuing to innovate and refine these technologies, VR can offer empowering and immersive experiences for individuals with limited mobility, contributing to their rehabilitation, well-being, and overall quality of life. That is the core reason for the work involved in this master thesis.

2.1.5 Categorization of 3D Interaction Techniques

3D interaction techniques encompass a wide range of methods and approaches for interacting with virtual environments. While the topic of 3D interaction techniques is complex, most interactions can be categorized into one of the following categories:

1. **Selection**: This category involves specifying which interactable object the user intends to interact with. Selection techniques enable users to identify and choose virtual objects within the environment for further actions or manipulation. Examples of selection techniques include pointing, ray-casting, or gaze-based selection.

2. **Manipulation**: Manipulation techniques allow users to modify the position, rotation, or scale of objects within the virtual environment. These techniques enable users to interact with and transform virtual objects, providing a sense of control and agency. Common manipulation techniques include grabbing and moving objects, rotating them, or scaling them to desired sizes.

3. **Navigation**: Navigation techniques facilitate user movement within the virtual environment. These techniques enable users to explore and tra-
verse the virtual space, often simulating natural locomotion. Navigation techniques can include walking or running in place, teleportation, flying, or using virtual vehicles to navigate the environment.

4. **System Control**: System control involves modifying settings or parameters of the virtual environment itself. This category of interaction techniques allows users to customize various aspects of the virtual environment, such as adjusting lighting, changing visual effects, or manipulating simulation parameters. However, for the scope of the thesis focused on enabling control for disabled individuals in VR, system control techniques, which are often specific to a particular environment or simulation, will not be addressed extensively.

Considering the specific goal of the thesis, which aims to empower disabled individuals to experience VR, the focus will primarily be on selection, manipulation, and navigation techniques. These categories of interaction techniques directly contribute to enabling users to interact with and navigate the virtual environment, ensuring an inclusive and accessible VR experience for individuals with disabilities.

### 2.1.5.1 Selection Techniques

In VE, the selection phase is a crucial aspect of user interaction. Whether it occurs before manipulation or navigation, the selection phase serves to disambiguate the user's intended targets for movement or manipulation. In traditional computer software, object selection is typically accomplished using virtual pointing, where the user selects an object in the direction pointed by a device, such as a mouse, which is well-suited for accurately pointing towards objects on a screen. In the context of VR, virtual pointing, also known as ray-casting, has gained significant popularity. However, the use of virtual hands, which allows users to interact and grab objects based on the position of their virtual hands, often mimicking their real-world hand movements, is also prevalent.

A comprehensive synthesis of various selection techniques can be found in the work of Arguelaguet [44]. This work provides valuable insights into the different methods employed for object selection in VE, aiding researchers and developers in selecting appropriate techniques based on their specific requirements and design goals.

Evaluation of selection techniques in VE is essential to ensure their effectiveness, efficiency, and user satisfaction. [45] have proposed guidelines
and a checklist for evaluating selection techniques. These guidelines offer a systematic approach to assessing the performance of selection techniques, considering factors such as accuracy, speed, user effort, and learnability. By following these guidelines, researchers, and practitioners can conduct comprehensive evaluations of selection techniques, facilitating informed decision-making and optimization of user interaction in VE.

In conclusion, the selection phase holds significant importance in virtual environments, serving to disambiguate user intentions for manipulation or navigation. Virtual pointing, such as ray-casting, and virtual hands are popular techniques used for object selection in VR. The work by Arguelaguet et al. provides a comprehensive overview of various selection techniques, aiding in the selection of appropriate methods for specific design requirements. Additionally, the evaluation guidelines and checklist proposed by Bergström et al. offer a systematic approach to assess the performance and effectiveness of selection techniques, enabling researchers and developers to optimize user interaction in VE.

2.1.5.2 Manipulation Techniques

In VR, manipulation plays a crucial role in interacting with the VE. It enables users to perform various actions such as moving objects in an industrial context or conducting surgical procedures in medical training simulations. Typically, manipulation in VR is achieved through the use of a real-world metaphor, where users can reach out, grab virtual objects, and manipulate them using body, arm, and hand motions.

However, a challenge arises when dealing with large objects in VR, for example [46]. Due to the physical limitations of the user’s space and the tracking technology used, it becomes difficult for users to interact with and accurately place large virtual objects. The user must be in close proximity to the object, which can be challenging when dealing with objects of substantial size or when physical movement is constrained.

Additionally, the requirement for users to physically move their bodies to manipulate objects in VR poses limitations. Some users may have mobility impairments or physical disabilities that restrict their ability to freely move their bodies. This can significantly impact their ability to interact with the virtual environment and perform manipulations effectively.

Considering these challenges, alternative approaches to manipulation in VR are being explored [47]. By advancing the field of manipulation techniques in VR, the goal is to enhance the user experience and enable a broader range
of users to effectively interact with and manipulate objects within virtual environments.

2.1.5.3 Navigation Techniques

Navigation is a fundamental aspect of interacting with VE allowing users to move and explore within these simulated spaces. The choice of navigation technique depends on the size and characteristics of the VE. In smaller VE, real-walk, a technique where the user physically moves in the real world, thereby causing their virtual counterpart to move synchronously, can be employed.

However, when dealing with larger VE, alternative techniques are required to enable efficient traversal through the virtual scene. A comprehensive list of major navigation techniques has been proposed by the Locomotion Vault project [48], offering valuable insights into the diverse approaches available.

According to [49], four main techniques can be identified:

• **Walking-Based**: In this technique, the user performs repetitive motor actions in the real world to navigate within the virtual world. Walking-based techniques often provide high-fidelity experiences, but they may pose challenges in terms of accessibility for individuals with mobility limitations.

• **Steering-Based**: In steering-based techniques, the user has continuous control over the direction of navigation. This category can be further subdivided into spatial steering techniques, where navigation is controlled using body-part movements or gestures, and physical steering techniques, which involve the use of physical props or controllers to direct navigation.

• **Selection-Based**: Selection-based techniques focus on where the user wants to go rather than how to get there. Two primary techniques within this category are target selection and route planning. In target selection, the user selects a specific destination, and the viewpoint is instantaneously moved to that location (e.g., teleportation). Route planning allows the user to preselect a desired route before initiating navigation. However, route planning techniques are less commonly implemented compared to target selection techniques.

• **Manipulation-Based**: Manipulation-based techniques involve the user manipulating their viewpoint or the virtual world itself through gestures.
These gestures can include rotation, translation, and scaling. One example of a manipulation-based technique is the World-In-Miniature concept, proposed by LaViola Jr. [50], where users interact with a miniature representation of the VE to control their viewpoint or navigate within the virtual space.

In summary, navigation techniques play a crucial role in enabling user movement within virtual environments. Depending on the size and requirements of the VE, different techniques can be employed. Walking-based techniques offer high-fidelity experiences, but may have accessibility limitations. Steering-based techniques provide continuous control over navigation, while selection-based techniques prioritize the selection of destinations. Manipulation-based techniques involve gesture-based interactions for viewpoint control. Understanding and utilizing these navigation techniques effectively contribute to enhancing user immersion and interaction in virtual environments. In our case, steering should give the user a better SoE considering they will be in control of the navigation at all time.

### 2.1.6 Eye-Tracking

Eye-tracking technology has emerged as a powerful tool that enables users to control simulations through the movement of their eyes. This technology has been commercially available for several decades on desktop platforms, and with the integration of eye-trackers into HMD, research and development in the field of eye-tracking has witnessed significant progress in recent years. The utilization of eye-tracking in simulations offers a novel and intuitive means of interaction, enhancing the user experience and opening up new possibilities for various applications. Presently, three primary techniques are employed for eye-tracking, each with its own unique advantages and implementation methods.

- The first technique, known as Electro-Oculography (EOG), involves the measurement of the eye’s orientation by utilizing electrodes placed around the eye to track its resting potential. This method allows for precise tracking of eye movements, enabling accurate control and manipulation within a simulation.

- The second technique, called Scleral Search Coils, operates by tracking the orientation of a wire loop embedded within a contact lens worn by the user. As the wire loop moves, it induces an electric current in Helmholtz coils positioned on the user’s head. By detecting and
analyzing the changes in the induced current, the eye’s movements
can be accurately captured and translated into simulation commands.
Scleral Search Coils offer a reliable and precise method of eye-tracking,
enabling seamless interaction with virtual environments.

- The third technique, Video-Oculography, relies on capturing images of
the user’s eye using cameras integrated within the VR headset. These
cameras record the movement and behavior of the eye, allowing for real-
time tracking and analysis. Video-Oculography offers a non-invasive
and user-friendly approach to eye-tracking, as it eliminates the need for
additional external equipment and relies solely on the built-in cameras
within the VR headset.

Currently, the primary application of eye-tracking technology in simu-
lations revolves around the analysis of design and experimental scenarios.
By tracking the user’s eye movements, researchers, and designers can gain
valuable insights into human behavior and preferences within simulated
environments. This data can then be used to refine and improve the design
of simulations, enhancing their usability and effectiveness.

In conclusion, the integration of eye-tracking technology into simulations
has revolutionized the way users interact with virtual environments. While
the main application of eye-tracking in simulations currently revolves around
design and experimental analysis, its potential for various other domains,
such as selection or navigation, is vast, promising further advancements and
widespread adoption in the future. It will be used in this master thesis, in order
to give control back to users who cannot use their whole body.

2.1.7 Evaluation

Evaluating the effectiveness and user experience of navigation and manipula-
tion in virtual reality (VR) systems requires the establishment of appropriate
evaluation metrics. These metrics serve as quantitative and qualitative
measures to assess the performance and usability of VR interactions.
Key evaluation metrics commonly employed in VR research include task
completion time, accuracy, error rates, and user satisfaction ratings [51, 45].
Task completion time provides insights into the efficiency of navigation
and manipulation techniques, allowing for comparisons between different
approaches. Accuracy metrics gauge the precision and correctness of
users’ interactions, providing an indication of the reliability of the system.
Error rates help identify and analyze user errors and misconceptions,
enabling iterative improvements. Furthermore, user satisfaction ratings capture the subjective experience and overall satisfaction of users, providing valuable feedback on the usability and enjoyment of the VR system. By utilizing a combination of these evaluation metrics, researchers can obtain a comprehensive understanding of the performance and user experience of navigation and manipulation in VR environments, enabling the refinement of existing techniques and the development of more effective interaction paradigms.

2.2 Related Work

This section provides an overview of the existing literature and research conducted on the specific subject addressed in this master thesis. It serves to contextualize the current study within the broader academic landscape and highlight the relevant advancements, theories, and findings that have been previously explored.

2.2.1 Manipulation of the Sense of Agency

Numerous studies have been conducted to explore the manipulation of the sense of agency (SOA) and embodiment in virtual reality (VR) environments. These investigations have aimed to better understand the factors that influence user experiences and interactions within VR simulations. In this section, we discuss several relevant studies that provide insights into the impact of different factors on the sense of agency and embodiment. We mostly study the sense of agency, considering it will be the one subcomponent of the sense of embodiment that should change the most when modifying control methods.

[52] investigated how user posture influences the impression of locomotion during VR observation. They found that the sense of agency was dependent on the user’s position, with the standing posture resulting in a stronger impression of locomotion. However, as our study focus on participants with quadriplegia, standing posture is not applicable, making our research particularly relevant for exploring alternative methods of interaction.

[53] examined the effect of a first-person perspective while being seated on the illusion of agency over a walking virtual body. They demonstrated that a first-person perspective from a seated position was sufficient to induce a strong illusion of agency over the virtual body, suggesting the potential for our experiment to elicit similar impressions.
[54] utilized brain-computer interface (BCI) paradigms to investigate the influence of different control methods on agency and responsibility over virtual movements. They found that having more control over the virtual arm, particularly through sensorimotor area activation, resulted in higher levels of agency and responsibility.

[55] developed a partial-visuomotor technique based on upper-body motion tracking to induce a sense of embodiment for individuals with reduced lower-body mobility. Despite a small number of participants, they observed positive responses toward the sense of embodiment using their design, aligning with our long-term objectives.

[56] investigated the relationship between locomotion techniques and embodiment in VR. While they did not find a significant impact of locomotion techniques on the sense of embodiment, it is crucial to note that all techniques provided the same level of avatar control, which motivates our focus on creating techniques with varying levels of control.

[57] explored the influence of latency, action modality, and display modality on the sense of agency in VR. Their findings indicated that voice commands led to weaker sense of agency, and higher latency reduced the explicit sense of agency. This highlights the importance of action modality in influencing the sense of embodiment in our study.

[58] investigated the impact of body ownership on illusory self-attribution of speaking in VR. Participants exhibited strong subjective sense of body ownership and agency over a synchronous virtual body. Although these results are intriguing, it is important to acknowledge that this study has not been widely reproduced.

[59] evaluated the effects of latency on perceptual judgments and motor performance in VR. They observed that even with a delay of up to 350ms, the sense of agency was not entirely disrupted, emphasizing the robustness of the sense of embodiment to certain delays.

[60] demonstrated that the type of observation of the virtual body influenced the sense of agency. This finding highlights the importance of the avatar’s presentation in shaping user experiences. In this regard, we decided to add mirrors in the simulation to help the users see their virtual body.

[61] investigated the sense of agency in human-machine interactions and found that the level of automation in a task influenced the sense of agency. While this study was not conducted in VR, it aligns with our understanding of agency and its relation to task control.

In summary, these related studies shed light on the factors that influence the sense of agency and embodiment in VR environments. By leveraging
these findings and building upon them, our research aims to contribute to the development of novel interaction techniques that enhance the sense of agency and embodiment for avatar control in VR.

2.2.2 Hands-free interaction

We aim to push the boundaries of “hands-free” interactions by exploring the concept of “body-free” interactions. Previous research, such as the systematic review conducted by P. Monteiro [62], has shown that previous research predominantly focused on using voice as an alternative interaction method. However, when considering future social situations, voice interaction may not be a suitable option. Additionally, most of the reviewed studies have not adequately evaluated the sense of embodiment or the viability of their interfaces, as highlighted by the following guidelines on object selection and manipulation evaluation [45]. Following Hornbaek’s categories [63] only 3.8% of them asked for immersion, like [37], emphasizing a gap in the literature regarding this topic.

[64] compared various input methods for user experience in VR and found that head-tracking was the most effective alternative, followed by eye-tracking. However, dwell and blinking were considered less favorable, although they served as reasonable alternatives when a button press was not possible.

[11] developed a system to compare different input modalities, including a combination of eye gaze and Electromyography (EMG) controller. Notably, the evaluation focused primarily on a pointing task in a 3D space. The throughput of EMG + gaze was superior to gaze with dwell, but lower than the throughput of head dwell. The estimated workload measured through the use of the NASA-TLX [15] indicated that the EMG + gaze workload was lower than other conditions that involved body movement.

BCI are also of interest in this research. [65] demonstrated the extensive use of BCI in VR, particularly in conjunction with eye-tracking technology.

While some interesting and unconventional interfaces, such as a mouth switch using a baby pacifier for interaction, have been explored, they are not relevant to our focus on using only eye or head movements [66].

Overall, the existing literature provides valuable insights into alternative interaction methods in VR, but there is a need to further explore the concept of “body-free” interactions and evaluate their impact on user experience and, hopefully, embodiment.
2.2.3  **Eye-Tracking for Selection**

In order to mitigate eye fatigue associated with prolonged use of eye tracking systems, researchers have proposed various techniques. [8] emphasized the potential of eye tracking in alleviating gorilla arm syndrome, a condition caused by extended use of manual input devices, such as mice and keyboards. However, they noted that gaze-based pointing methods often suffer from errors, noise, and a significant challenge known as the Midas touch problem [19]. The Midas touch problem refers to the issue of unintentionally activating elements on the screen simply by looking at them. Consequently, researchers have explored numerous strategies to address this challenge and prevent inadvertent activations.

Although several techniques have been proposed to mitigate eye fatigue, none of them have been specifically utilized in the context of avatar control, and none have been evaluated in terms of embodiment. In the literature, two distinct types of research papers can be identified based on their focus: (1) those comparing eye-based interaction with alternative input methods, and (2) those introducing novel eye-based interaction techniques.

2.2.3.1  **Comparing Eye-Based Interaction with Alternative Input Methods**

Several studies have compared eye-based interaction with other input methods to evaluate their performance and suitability. [67] conducted a study comparing head, gaze, mouse click, and dwell interaction. The study found almost identical throughput for both head and gaze pointing, indicating that these methods could be interchangeable. [68] investigated the performance of eye + button interaction compared to head + button interaction. The results showed that the head + button interaction yielded significantly better error rates, selection times, and throughput compared to the eye + button interaction. [69] examined the effectiveness of eye aiming versus hand controllers for an aiming task. The study revealed that while gaze aiming resulted in lower perceived cognitive load, the accuracy was lower compared to using hand controllers, despite both methods having similar task durations. Further research in this direction is necessary to explore the trade-offs between different input methods in avatar control.
2.2.3.2 Introducing Novel Eye-Based Interaction Techniques

Researchers have also proposed novel techniques that leverage eye tracking to enhance interaction and overcome specific challenges.

[70] introduced the Eye & Head technique, which combines head and eye movements for selection. Selection occurs when both the head and the eye are fixated on the same object. This technique exploits the synergy between eye-tracking and head-tracking, presenting an intriguing approach for avatar control. In another work by Sidenmark, the Outline Pursuit technique was proposed to address the occlusion problem. This method tracks the pursuit eye movement toward a moving stimulus, allowing users to select objects even when they are partially occluded [26].

[25] presented the DualGaze technique, which addresses the Midas touch problem by introducing a two-step selection process. Users pass their gaze over an object, and a flag appears next to the object location. To select the object, users must navigate to the flag, mitigating the risk of accidental activations. However, in our case, this technique was found to be less effective than dwell interaction and was thus excluded from further consideration.

[71] proposed several techniques, including Duo-Reticles, Radial Pursuit, and Nod and Roll. Duo-Reticles involves the use of an “inertial reticle” that moves towards the eye-gaze reticle but never intersects with it. To make a selection, users must redirect their gaze back to the inertial reticle. Radial Pursuit deals with cluttered objects and employs a sphere surrounding the target object. Objects within the sphere are evenly distributed, and users can pursue an object for selection based on distance and a given confidence threshold. Nod and Roll utilizes the eyes’ ability to stabilize on an object despite head movements. Users can select an object using their eyes and confirm the selection by nodding their head.

[72] explored the use of wink and double blink gestures for controlling a virtual reality (VR) game similar to Subway Surfer. However, this study did not provide a direct comparison with other techniques, offering only a “Percentage of VR control” metric.

[73] investigated eye winks, electromyography (EMG), and speech-based interaction for controlling robotic devices. The study found that eye-tracking achieved the lowest mean error rate, further supporting the potential of eye-based interaction in avatar control.

While additional techniques such as Gaze and Pinch [74] exist, these methods involve body movements and were excluded from consideration due to our focus on non-body-based interaction techniques.
Despite the non-use of these techniques in the final version of this work, we tried to implement Duo-reticle in the selection task but after trying it in a pilot study, the method was rejected with the reasons that dwell was both easier to understand and quicker to use with no prior experience. Thus, we present them and know they are available despite not being used in the last version of this work.

### 2.2.4 Eye-Tracking for Manipulation

Manipulation tasks using eye-tracking are less common and more complex, resulting in a smaller number of research papers. Nevertheless, three papers provided valuable insights into this area.

[27] implemented three rotation techniques based on the number of rotation axes. The RotBar technique created a bar for each rotation axis (X, Y, Z), the RotPlane technique employed orthogonal planes, and the RotBall technique utilized an arcball for user-perspective roll manipulations. However, the study did not include a baseline comparison with real controllers. The fastest rotation was achieved in a single-axis task, yet even the fastest completion time was 39 seconds, indicating the difficulty of achieving the desired rotation.

[75] introduced four techniques for interacting with objects at different distances. This study employed a combination of eye and hand movements to control manipulation. The eyes were primarily used for object selection and translation within a limited spatial region, while most of the manipulation was performed using the hands. The manipulation time for all four techniques was similar, averaging around 12 seconds.

[13] approach decoupled the use of both eyes and employed the closure of one eye to trigger different actions depending on the state of the left or right eye. Their study primarily focused on 2D tasks, such as navigating a user interface. The findings revealed that using both eyes and a manual trigger outperformed using the closure of one eye in all aspects, although the approach utilizing eye closure was intriguing.

These papers contribute valuable insights into eye-based manipulation techniques. Considering the difficulty of giving a good user experience for rotation using only the eyes, we decided not to include rotation in our manipulation task, deeming it too challenging for an exploratory study. In the subsequent sections, we will present our own approach to eye-based manipulation and evaluate its effectiveness in terms of task performance and user experience.
2.2.5 Eye-Tracking for Navigation

Navigation tasks using eye-tracking are quite common and have been explored, since navigation is an important part of every VR experience.

[76] directly compared three navigation techniques: Gaze-Directed (forwards is the direction the user faces), Pointing (the direction is chosen by pointing with a controller) and Teleport. Participants had to reach their destination while collecting object on the route. It appeared that teleport was better for cybersickness and performance but led to an increased chance of missing tokens on the road, making it a bit less useful for exploration task.

[12] compared three different triggers techniques for teleportation: Eye-wink, a mouth gesture and dwell. The baseline was a controller-based teleportation. An interesting point to note is the fact that they excluded voice as a trigger because of social situations. In general, dwell and wink were the most robust method for errors and accuracy, but dwell was considered slow and the best method for speed were wink and controller-based. Overall, head and eye method were good alternative to controller methods.

[77] introduced discrete and continuous input for navigation through an UI. In a discrete input, you can switch the navigation on and off by looking at it, while for continuous input, you have to look at it for it to be active. They also proposed continuous or gradient-based method. In a continuous situation, the input is whether on or off while in a gradient based method, the velocity of the input is gradually increased or decreased depending on where you look. Their finding seem to privileged continuous x gradient based method for two main reasons being that the discrete activation needed a dwell time, making it slower and the speed was also a point of reproach because it was deemed to be too slow for long travel and too fast for small travel. However, considering the fact that they were not in VR, they did not have to account for potential cybersickness that could arise with the use of a varying speed. Overall, this comforts us in the idea that steering with UI is possible.

[20] investigated the efficiency of seven different interaction techniques involving eye and head tracking for navigation in VR environments. The goal was to understand how the combination of eye and head movements impacts performance compared to head-only interactions. Despite some challenges related to eye-tracking issues and imprecision, the results indicated that head + eye interactions were slightly less efficient than head-only interactions, though the difference was minimal. Interestingly, the authors suggested that the use of a better eye-tracker could potentially enhance eye-tracking performance to match that of head-tracking. This conclusion highlights the need for further
research, such as our current study, to explore the potential improvements in performance achievable through better eye-tracking technology.

2.3 Summary

This scientific summary provides an overview of existing literature and research related to the sense of agency manipulation, hands-free and body-free interactions in VR, eye-tracking for selection, manipulation, and navigation in VR environments.

- Manipulation of the Sense of Agency: Studies have explored how factors like user posture, virtual exercise, brain-computer interface, and locomotion techniques influence the sense of agency and embodiment in VR. The research aims to understand how different interactions affect the user’s perception of control and agency over virtual experiences.

- Hands-free Interaction: The focus is on exploring “body-free” interactions in VR, going beyond voice-based interactions. Various input modalities, such as head and eye tracking, EMG controllers, and brain-computer interfaces, have been investigated. The research aims to develop novel techniques for hands-free interactions and evaluate their impact on embodiment and user experience.

- Eye-Tracking for Selection: Studies have compared eye-based interaction with other input methods, introduced novel eye-based techniques, and investigated eye fatigue mitigation strategies. Researchers have proposed different eye-based selection techniques, such as Duo-Reticles and Radial Pursuit, but further research is needed to explore trade-offs between different input methods.

- Eye-Tracking for Manipulation: Research on eye-based manipulation is less common but offers insights into rotation and translation techniques using eye tracking combined with other modalities like hand movements. The studies highlight the complexity of manipulation tasks with eye tracking and the need for further exploration in this area.

- Eye-Tracking for Navigation: Studies have compared eye and head tracking interactions for navigation in VR. Results have shown that eye and head interactions can be efficient for navigation, although challenges related to eye-tracking accuracy should be addressed.
Overall, the literature provides valuable insights into the manipulation of the sense of agency and thus the manipulation of the sense of embodiment, the development of hands-free interactions, and the potential of eye-tracking for selection, manipulation, and navigation in VR environments. However, further research is necessary to explore the potential improvements in eye-tracking technology and its impact on user experiences in VR.
32 | Background
Chapter 3

Methods

This chapter outlines the methodologies and experimental designs adopted to develop and validate the hypotheses presented in this study. An overview of the whole process is presented in figure 3.1.

3.1 Research Process

The research process was divided into five distinct phases to systematically address the objectives of this study. Each phase contributed to the development and refinement of the research methodology, ensuring a comprehensive investigation of avatar control techniques in virtual reality (VR).

Phase 1: Literature Review and Gap Analysis The initial phase involved an extensive literature review to gather relevant references and analyze existing research gaps. This step provided an in-depth understanding of the current state of avatar control in VR, identifying areas that had not been adequately explored. This knowledge formed the foundation for defining the research objectives and shaping subsequent phases.

Phase 2: Technique Exploration and Selection In the second phase, a diverse range of techniques for object selection, translation, and rotation were explored. Multiple approaches were implemented and evaluated. This phase aimed to experiment with various methods to ensure comprehensive coverage and to identify techniques that exhibited promising outcomes. A preliminary testing phase involving a small group of testers was conducted to assess the feasibility and effectiveness of the implemented techniques.

Phase 3: Refinement and Pilot Testing The third phase focused on refining the application and techniques based on feedback from the
preliminary testing. Pilot users were engaged to participate in more extensive experiments. Their experiences and insights were gathered to further optimize the techniques and application design. The pilot users tried different conditions. The first iteration was a test with selection techniques, comparing dwell and Dualgaze for selection with the eyes. Dualgaze was discarded because all pilot users judged it to be harder to learn, less intuitive, and the overall performance was not better. The second iteration was with the speed of the avatar’s hands and walk speed, as well as the placement of the rotation facilitator. This iterative process ensured that the research tools were well-tailored for the user study in the subsequent phase.

**Phase 4: User Testing and Evaluation** In the fourth phase, the refined techniques were tested with a larger group of users. This phase involved rigorous user testing to gather quantitative and qualitative data on the performance and user experience of the different avatar control techniques. Participants’ interactions were systematically recorded, and physiological measures were collected to provide a comprehensive understanding of the user interactions and responses.

**Phase 5: Data Analysis and Conclusion** The final phase focused on analyzing the accumulated data from the user testing phase. The collected data underwent thorough statistical analysis and qualitative interpretation to draw meaningful insights. The results were compared with the predefined hypotheses to determine the effectiveness and suitability of the implemented techniques. This analysis also allowed for the formulation of conclusions and recommendations based on the research findings.

### 3.2 Experimental Setup

The primary aim of this research paper was to perform a comprehensive comparison of distinct techniques for avatar control, with a focus on evaluating their impact on the Sense of Embodiment (SOE), User Experience (UX), and performance. The experiment encompassed the following two tasks:

- **Pick & Place Task:** In this task, participants were required to precisely align the position of an object with a predefined target location. The task involved both selecting the object and accurately translating it to match the target’s position.

- **Navigation Task:** The navigation task challenged participants to navigate within a virtual environment, collecting glowing cubes placed throughout a room.
There is a difference to make between the input modality and the conditions for the two tasks, as outlined in the figure 3.2.

Figure 3.2: The distinction between input modality and the methods used for the selection & manipulation task and the navigation task.

When starting the experiment, the participants were attributed an avatar, depending on how they identified. The two avatars used can be seen in fig 3.3. The avatars were taken from the commonly used “Microsoft Rocketbox Avatar Library” [78].
Figure 3.3: The two avatars used for both tasks. Participants were attributed one at the beginning of the experiment and would keep the same during the whole experiment.

### 3.2.1 Pick and Place

This task includes both selection and manipulation aspects, constituting two of the most prevalent activities within VR user interaction scenarios. The chosen tasks mirror common real-world interactions, contributing to the ecological validity of the experiment.
To simulate a context closely resembling real-world VR experiences and to enhance user engagement, a virtual kitchen environment was designed. Participants engaged in assembling virtual burgers, imbuing the experiment with a sense of purpose. This approach not only facilitated the relevance of the results, but also alleviated potential boredom, consequently augmenting the overall user experience. An image of the task environment, seen from a
first and third perspective, can be found in figure 3.4

The burger assembly task entailed picking and placing various elements, creating a seamless progression of burger components: Bread / Steak & Cheese / Tomatoes / Crudités / Bread. Elements were visually distinguishable but consistently positioned, and maintained the same size. Each successfully placed element enriched the user’s burger, instilling a sense of accomplishment. In addition, auditory cues were integrated, accompanied by post-processing enhancements, engendering a more immersive VR experience.

To ensure practicality and feasibility, the user’s hand posture altered when selecting objects, mitigating physically impossible scenarios. Given the study’s focus on enabling disabled individuals to control avatars, the selected factors directly influenced the avatar control: the degree of control and input modality.

Selection techniques like dwell, dual gaze, and wink were excluded from investigation, as they were unlikely to significantly affect embodiment results based on the agency definition. Consequently, the “dwell” technique was adopted for object selection, consistent with prior research.

For object manipulation, a straightforward approach was employed. A ray projected from the midpoint of the user’s eyes determined the selected object’s destination. When an object was hit, it moved to the point of contact; in the absence of contact, it moved a fixed distance. Notably, rotation techniques were omitted due to their complexity and lack of significant results in prior research in this domain.

A run consisted of picking and placing one element. In order to complete a condition, the participant had to go through 5 runs sequentially.

An example video of the task and the four conditions can be found here: [YouTube - Interaction](#). Be careful that the video does not render the environment as it would in a VR headset (notably the field of view).

### 3.2.1.1 Degree of Control

The degree of control refers to the extent of control participants exert over the avatar. Two distinct levels were implemented:

- **Full Hand Control:** Participants selected a hand and moved it, akin to real-world hand movements. For object interaction, users designated a hand and maneuvered it to the object’s location, possessing precise control over the hand’s movement. It will be referred to as the “avatar” target.
• **Object Control:** Participants selected an object, prompting the nearest hand to automatically engage with and manipulate the object. Subsequently, users could reposition the object. It will be referred to as the “object” target.

Advanced controls, such as individual finger manipulation and hand rotation, were omitted due to their intricacy and potential cognitive burden. Notably, additional control levels, including automated or triggered animations, were excluded to maintain research focus.

Note that Full Hand Control entailed two additional steps compared to Object Control, potentially diminishing its anticipated usability but elevating its perceived agency.

**3.2.1.2 Input Modality**

Given the research’s target audience, that might or might not be able to use their neck, a comparison of two input modalities was interesting. Two chosen input modalities were assessed:

- **Eye-Tracking:** Selection and manipulation aligned with eye gaze. Notably, eye gaze altered with head rotation, rendering it a dual-input technique.

- **Head-Tracking:** Selection and manipulation synchronized with headset orientation.

To prevent participants from comparing these modalities to controllers, which utilize tangible body motion, controller inputs were intentionally excluded. Controllers’ established usability could confound participants’ assessments of Eye-Tracking and Head-Tracking, thereby reducing the discernible contrast between the two techniques.

**3.2.2 Navigation**

Navigation constitutes a pivotal aspect of any experience extending beyond a fixed location. Particularly for disabled individuals, unrestricted movement is often unattainable. This task aimed to empower disabled participants by providing them with the means to navigate their environment autonomously.
The navigation task required participants to traverse a series of designated points, represented by glowing cubes for visibility. To simulate real-world scenarios and enhance applicability, a fast-food setting was chosen. Participants navigated within this setting, replicating customer service tasks to foster a more authentic engagement. An image of the task environment can be found in figure 3.5.

Navigation destinations were determined using ray-casting, analogous to the method employed in the pick & place task. The navigation was initiated upon trigger activation. Since disabled participants might not have the capability to physically move, rotating their chairs was disabled. Instead,
camera rotation was facilitated by directing the gaze toward the screen edges. There were two outlined squares on the left and right side of the field of view, and when looking at one, the rotation in the corresponding sense would occur.

In order to prevent cybersickness, a reduction of the field of view was instantiated with a vignette effect every time the user was in rotation or navigation.

Uniform parameters were maintained for speed and rotation across all runs. Navigation initiation was triggered by a prolonged blink to circumvent the Midas Touch Problem inherent in dwell and dual gaze techniques.

A run consisted of going from one cube to another cube. In order to complete a condition, the participant had to go through 4 runs sequentially.

The cubes were positioned in a square layout, and to prevent any learning pattern, they appeared alternately in both clockwise and counter-clockwise orders.

An example video of the task and the four conditions can be found here: [Youtube - Navigation](#). Be careful that the video does not render the environment as it would in a VR headset (notably the field of view).

### 3.2.2.1 Navigation Methods

To discern variations between free exploration approaches, two navigation methods were selected, mirroring the pick & place task:

- **Navigation Steering:** Participants entered a mode where their avatar moved toward the chosen destination in real time while the navigation steering was active. This allowed users to meticulously chart their path to each point, as the destination updated per frame. It will be referred to as the “Avatar” target.

- **Navigation Trigger:** Users initiated movement towards their destination, traversing the path until they reached the designated point. It will be referred to as the “Goal” target.

### 3.2.2.2 Input Modality

Consistent with the pick & place task, both eye-tracking and head-tracking modalities were examined.
3.3 Data Collection

Data collection was executed through user testing, with participants providing informed consent before engaging in the experiment. A within-subjects comparison design was employed, ensuring that each participant underwent all conditions (2 tasks and 4 conditions per task). Due to potential visuo-vestibular conflict-induced cyber-sickness associated with navigation, participants commenced with the completion of the 4 pick & place conditions prior to addressing the 4 navigation conditions.

To mitigate the influence of task learning on our outcomes, a balanced Latin-square design was implemented for both tasks.

3.3.1 Sampling

Considering that the target population for this study is disabled individuals, but due to medical constraints, we conducted the testing on a general healthy population. With each participant engaging in all conditions, the study employed a repeated-measures design.

To analyze the data, a repeated-measure 2-way ANOVA was chosen due to the presence of two conditions (input and method). The calculation of the minimum required sample size a priori was based on the following parameters:

- Utilizing Cohen’s effect size definition, a medium effect between methods was assumed, yielding an effect size of 0.5.

- The significance level (alpha error probability) was set to the standard 0.05.

- A power of 0.9 was considered.

- The study comprised four groups (conditions).

- Each participant contributed four measurements.

Based on these parameters, a minimum sample size of 12 was estimated. However, to enhance the statistical power, a total of 24 participants were recruited, providing a more robust basis for deriving meaningful results.
3.4 Experimental design and Planned Measurements

As emphasized in [45], it is advisable to prioritize either speed or accuracy as the primary dependent variable, rather than attempting to optimize both simultaneously. In alignment with this principle, our research focuses on speed as the key parameter of interest, given that, in most real-world applications of virtual reality (VR), it holds greater significance than accuracy about performance. Throughout each trial, excluding the tutorial phase, we recorded the time taken by participants to complete the tasks.

Upon arrival, participants commenced the experiment by providing their informed consent. Then, they filled a short demographic form. The initial steps involved selecting the correct avatar depending on the gender, adjusting the avatar to the appropriate scale and performing eye-tracking calibration. Subsequently, participants were exposed to specific combinations of independent variables and were afforded the opportunity to engage in practice runs to familiarize themselves with the task requirements, acting as a tutorial phase. Upon completion of the tutorial, participants were tasked with executing the assigned objective with each condition. Following this task performance, participants were prompted to complete the VEQ, available at A.1, to gauge their SoE. Additionally, participants were asked to complete the SUS questionnaire, available at A.2, evaluating the usability of the given condition, and the NASA-TLX questionnaire, available at A.3, designed to ascertain the perceived workload associated with each combination. Notably, the form completion process was conducted outside the VR environment to allow participants a break, as the experiment extended over a span of one hour. Upon returning to the next condition, recalibration of the eye-tracker was performed. Participants were then introduced to a different combination and proceeded to experience each condition for both tasks. Upon completing all four conditions for task 1, participants transitioned to task 2.

3.4.1 Hardware/Software used

The experiment was conducted utilizing a computer equipped with high-performance hardware, including an RTX 3090 graphics card, an Intel Xeon processor, and 32 GB of RAM. The experimental setup also incorporated a Vive Pro Eye headset, chosen for its integrated eye-tracking capabilities.

In terms of software infrastructure:
• Unity served as the primary development platform for crafting the experimental environment. The Inverse Kinematics (IK) of the avatar for the animation of the movements was managed using the FinalIK plugin.

• Audacity software facilitated the editing and segmentation of audio files.

• Blender was employed for the creation and modeling of the virtual environment.

• Data analysis was carried out using Python programming language.

• G*Power software was utilized for pre-experiment sample size calculation.

• Google Forms served as the tool for collecting form-based data.

• Evento Renater played a crucial role in session management and scheduling for the experiment.

3.5 Assessing reliability and validity of the data collected

Assessing reliability and validity is crucial in research to ensure the quality and accuracy of the data collected. In this section, we will discuss the measures taken to assess the reliability and validity of the data collected in our study.

3.5.1 Reliability Assessment

Reliability refers to the consistency and stability of the measurements taken. In our study, we ensured reliability through the following means.

Inter-Rater Reliability: In cases where subjective assessments were involved, such as the assessment of user experience and embodiment perception, multiple raters were involved in the analysis.

3.5.2 Validity Assessment

Validity refers to the extent to which the research measures what it intends to measure. In our study, we addressed different aspects of validity:
1. **Content Validity:** To ensure that our tasks and measures were representative of the constructs they aimed to assess, we consulted experts in the field of human-computer interaction and virtual reality. Their feedback helped us refine the tasks and the survey questions.

2. **Construct Validity:** We employed established scales and measures for constructs like user experience and embodiment perception. These scales have been widely used and validated in previous research.

3. **Face Validity:** Before conducting the main experiment, we conducted a pilot study with a small group of participants. Their feedback on the tasks, instructions, and survey questions helped us identify any ambiguities or potential issues with face validity.

4. **Concurrent Validity:** We compared the results of our study with existing research in the field to ensure that our findings aligned with the current understanding of the concepts being studied.

5. **Ecological Validity:** We designed the tasks and the virtual environment to simulate real-world scenarios as closely as possible. This helped enhance the ecological validity of our study, ensuring that the results could be generalized to similar real-world situations.

By addressing these reliability and validity concerns, we aimed to ensure that the data collected in our study is robust, accurate, and meaningful. This, in turn, strengthens the credibility of our research findings.

### 3.5.3 Validity of method

Assessing the validity of the chosen research method is essential to ensure that the approach used in the study is appropriate for addressing the research questions and achieving the objectives.

The method chosen for this study aligns closely with our research questions and objectives. Our aim was to compare different methods of avatar control based on the sense of embodiment, user experience, and performance. The experimental design allowed us to directly manipulate the variables of interest while controlling for potential confounding factors.

#### 3.5.3.1 Theoretical Foundation

The chosen method is rooted in well-established theoretical frameworks within the fields of human-computer interaction and virtual reality. The utilization of
established constructs such as sense of embodiment and user experience, along with validated measurement scales, provides a solid theoretical foundation for our research.

### 3.5.3.2 Pilot Testing

Prior to conducting the main experiment, we conducted a pilot study with a small group of participants. This pilot testing phase allowed us to identify any shortcomings in the experimental design, instructions, or procedures. The feedback from the pilot participants contributed to refining the methodology, enhancing its validity for the main study.

### 3.5.3.3 Expert Review

We sought the expertise of researchers and practitioners in the field of virtual reality and human-computer interaction to review our experimental design. Their input helped ensure that the methods used were sound and aligned with industry best practices.

### 3.5.3.4 Controlled Environment

To enhance the internal validity of our study, we conducted the experiments in a controlled virtual environment. This control over environmental variables minimized potential confounds that could affect the results and allowed for a more precise examination of the variables of interest.

### 3.5.3.5 Statistical Analysis

In analyzing the collected data, we employed robust statistical techniques, such as repeated-measures ANOVA. These analyses help establish the relationships between variables and support the validity of our findings.

Overall, the chosen method for our study demonstrates a high degree of validity, as it is well-founded in theory, aligned with research objectives, refined through pilot testing and expert review, and conducted in a controlled environment. By addressing these aspects, we ensure that the results derived from this method are trustworthy and contribute meaningfully to the field.
3.6 Data Analysis

In the data analysis phase, our primary tools were Python’s Pandas and Pingouin libraries, which enabled us to efficiently manage and process the collected data. For data visualization, we utilized Matplotlib and Seaborn, which facilitated the creation of insightful visual representations.

The primary statistical analysis involved the application of a two-way repeated measures ANOVA. This analytical approach was chosen to determine significant differences among the results obtained from different experimental conditions. Given that we incorporated only two conditions for each independent variable, we opted against using a t-test. This choice was informed by the fact that a significant outcome in the ANOVA inherently implies the superiority of one condition over the other in terms of the measured performance metric.
Figure 3.1: The timeline that was followed in this master thesis, presenting what was implemented and what was done, for the development, for the user testing and for the research process in general.
3.7 Conclusion

In this section, we outlined our research process, experimental setup, data collection methods, and the validity of our approach. We conducted a systematic investigation of avatar control techniques in virtual reality (VR) through five distinct phases. These phases culminated in user testing across two tasks, collecting data on speed, user experience, and embodiment perception. Our methodology demonstrated strong reliability and validity, rooted in established theories and expert review. Despite its strengths, this method has limitations, including a focus on low-level tasks and a healthy participant sample. Nonetheless, it provides a robust foundation for our study’s findings in the subsequent chapters.
Chapter 4

Results and Analysis

In this chapter, we present the results of our study and provide an analysis of the findings.

Symbol | p-value
---|---
* | < 0.05
** | < 0.005
*** | < 0.0005

Table 4.1: In every graph, a “*” between two conditions represents the p-value. If no star is present, then the results are not significant.

4.1 Results

Two participants were excluded from the results due to time constraints; they did not have the opportunity to participate in the navigation task after completing the selection and manipulation tasks.

4.1.1 Selection & Manipulation

The selection and manipulation task always preceded the navigation task.

4.1.1.1 Performance

For performance, specifically regarding speed, we made the decision to exclude data from the first repetition. The initial timing was unclear for some participants, resulting in disproportionate times, as can be seen on Figure 4.1. This exclusion was validated using a Geisser-Greenhouse correction,
yielding a p-value of 2.4e-8. Subsequent Tukey pair tests indicated significant
differences between the first run and all subsequent runs.

Figure 4.1: The time it took to complete a run, on average for each participant,
depending on whether it was the first, second, third, fourth, or fifth run of the
condition for the interaction task. The first run was systematically excluded
considering the time it took to complete was way higher than the rest of the
run, probably due to an unclear starting signal. For the signification of “*” see
4.1.

Regarding speed, Figure 4.2 displays the distribution, which mostly
follows a normal distribution, and the mean. We identified significant effects
for both the target and the method used, as shown in Table 4.2. Having control
only when selecting an object increased speed performance by 35% compared
to selecting the avatar. Additionally, using eye-tracking in addition to head-
tracking improved speed performance by 11%.
Figure 4.2: The time it took on average to complete a run for the interaction task (taking an ingredient, putting it in the plate) depending on the two independent variables, Degree of Control (see 3.2.1.1) and Input Modality (see 3.2.1.2). For the signification of “*” see 4.1.

Table 4.2: p-values for interaction performance (time taken for a run) based on the target and method used.

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>ddf1</th>
<th>ddf2</th>
<th>MS</th>
<th>F</th>
<th>p-unc</th>
<th>p-GG-corr</th>
<th>ng2</th>
<th>eps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target</td>
<td>$3.5 \times 10^4$</td>
<td>1.0</td>
<td>$2.3 \times 10^3$</td>
<td>$3.5 \times 10^7$</td>
<td>$1.6 \times 10^4$</td>
<td>$7.9 \times 10^{-12}$</td>
<td>$7.9 \times 10^{-12}$</td>
<td>$3.2 \times 10^{-1}$</td>
<td>1.0</td>
</tr>
<tr>
<td>Method</td>
<td>4.2</td>
<td>1.0</td>
<td>4.2</td>
<td>5.5</td>
<td>$2.8 \times 10^{-2}$</td>
<td>$2.8 \times 10^{-2}$</td>
<td>$5.4 \times 10^{-2}$</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Target * Method</td>
<td>$9.0 \times 10^{-1}$</td>
<td>1.0</td>
<td>$2.3 \times 10^3$</td>
<td>$9.0 \times 10^{-1}$</td>
<td>2.4</td>
<td>$1.4 \times 10^{-1}$</td>
<td>$1.4 \times 10^{-1}$</td>
<td>$1.2 \times 10^{-2}$</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Additional graphs can be found in Appendix B.1.1.

4.1.1.2 Sense of Embodiment

No significant effect was found on the overall Sense of Embodiment or any of its subparts (Acceptance, Control, and Change) for the interaction task. Further details are available in Appendix B.2.1.
4.1.3 System Usability Scale

No significant effect was found on the System Usability Scale for the interaction task. More information is provided in Appendix B.3.1.

4.1.4 Nasa Task-Load-Index

No significant effect was found on the Nasa Task-Load-Index for the interaction task. Additional data is available in Appendix B.4.1.

4.1.2 Navigation

The navigation task always followed the interaction task.

4.1.2.1 Performance

In contrast to the interaction task, the navigation task had a clearer starting point. The first run seemed similar to subsequent runs, and statistical analysis supported this observation with a p-value of 0.64, indicating no significant difference between the first and subsequent runs. Similarly, for clockwise and counterclockwise navigation, a p-value of 0.9 was found, indicating no significant difference.

Regarding speed, Figure 4.3 depicts the distribution, which largely follows a normal distribution, and the mean. We identified significant effects for both the target and the method used, as shown in Table 4.3. Maintaining control off the path at all times and navigating without selecting a destination improved speed performance by 36% compared to selecting the destination. Additionally, using head-tracking instead of eye-tracking enhanced speed performance by 12%.
Results and Analysis

Figure 4.3: The time it took on average to complete a run in the navigation task (going from one glowing cube to another glowing cube) depending on the two independent variables, Navigation Method (see 3.2.2.1) and Input Modality (see 3.2.1.2). For the signification of “*” see 4.1.

Table 4.3: p-values for navigation performance based on the target and method used.

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>ddf1</th>
<th>ddf2</th>
<th>MS</th>
<th>F</th>
<th>p-unc</th>
<th>p-GG-corr</th>
<th>ng2</th>
<th>eps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target</td>
<td>$7.9 \times 10^4$</td>
<td>1.0</td>
<td>2.2</td>
<td>$7.9 \times 10^4$</td>
<td>$5.6 \times 10^4$</td>
<td>1.7</td>
<td>$1.7 \times 10^{-7}$</td>
<td>$1.8 \times 10^{-1}$</td>
<td>1.0</td>
</tr>
<tr>
<td>Method</td>
<td>$1.1 \times 10^2$</td>
<td>1.0</td>
<td>2.2</td>
<td>$1.1 \times 10^2$</td>
<td>5.0</td>
<td>$3.5 \times 10^{-4}$</td>
<td>$3.5 \times 10^{-2}$</td>
<td>$3.0 \times 10^{-4}$</td>
<td>1.0</td>
</tr>
<tr>
<td>Target * Method</td>
<td>$3.3 \times 10^{-2}$</td>
<td>1.0</td>
<td>2.2</td>
<td>$3.3 \times 10^{-2}$</td>
<td>$9.8 \times 10^{-4}$</td>
<td>$9.8 \times 10^{-1}$</td>
<td>$9.8 \times 10^{-1}$</td>
<td>$9.0 \times 10^{-4}$</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Additional graphs can be found in Appendix B.1.2.

4.1.2.2 Sense of Embodiment

A significant effect was found for the control (agency) measure depending on the input modality, with a p-value of 2.8e-2. Figure 4.4 illustrates the results, indicating a 10% increase in agency when using head-tracking instead of eye-tracking.
Figure 4.4: The sense of agency score in the navigation task, referred to as the control score here, one of the subcomponent of the SoE, measured using the VEQ and displayed depending on the Input Modality. For the signification of “*” see 4.1.

Additional graphs can be found in Appendix B.2.2.

4.1.2.3 System Usability Scale

A significant effect was found for the System Usability Scale measure depending on the input modality, with a p-value of 1.4e-3. Figure 4.5 illustrates the results, indicating an 18% increase in usability when using head-tracking instead of eye-tracking.
Figure 4.5: System Usability Scale (SUS) score depending on the Input Modality for the navigation task. For the signification of "*" see 4.1.

Additional graphs can be found in Appendix B.3.2.

4.1.2.4 Nasa Task-Load-Index

A significant effect was found for the Nasa Task-Load-Index measure depending on the input modality, with a p-value of 3.5e-4. Figure 4.6 illustrates the results, indicating a 39% increase in task load when using eye-tracking instead of head-tracking.
Figure 4.6: Nasa Task-Load-Index (NASA-TLX) Score depending on the Input Modality for the navigation task. For the signification of "***" see 4.1.

Additional graphs can be found in Appendix B.4.2.

4.2 Impact of Virtual Reality and Video game Habit

Before starting the experiment, all participants completed a questionnaire assessing their familiarity with both VR and video games on a scale from 0 to 3, as can be found in Appendix A.4. We observed that lacking experience in VR had an impact on performance in both interaction and navigation, although the effect was relatively small, as seen in Figure 4.7 for interaction and Figure 4.8 for navigation. Pairwise comparisons using Tukey tests confirmed significant differences between different experience levels, as shown in Table B.1 and Table B.2 for interaction and Table B.3 and Table B.4 for navigation. In these tables, "A" and "B" represent different experience levels, and the table compares run durations.
Figure 4.7: Time taken for a run in the selection & manipulation task, depending on VR and Video game Experience. The VR experience seems to be playing a role in the results, considering that people who never tried VR have worse results than others. For the signification of "*" see 4.1.
In conclusion, this chapter has provided a comprehensive overview of the results and their significance in our study. These findings shed light on the effectiveness of different input modalities and control methods in VR, particularly for individuals with limited mobility. While no definitive answers have emerged, these results serve as a foundation for future research and development in the pursuit of more inclusive and user-friendly VR experiences.

Figure 4.8: Time taken for a run in the navigation task, depending on VR and Video game Experience. For the signification of "*" see 4.1.

Additional tables can be found in Appendices B.5.1 and B.5.2.

4.3 Conclusion
Chapter 5

Discussion

This chapter provides a comprehensive discussion and situates this work in a broader context. The research question 3, is discussed inside both 5.1 and 5.2.

5.1 Research Question 1: Selection and Manipulation

This section of the discussion focuses on the results related to Research Question 1 1.2.1.1, which examines the impact of different methods of eye-tracking on the sense of embodiment (SoE) and usability in virtual reality. Research Question 1 aims to understand how different input modality and control methods during an interaction task influences performance, usability, and the sense of embodiment.

5.1.1 Performance Improvement with Eye-Tracking

In the context of our interaction task, it is noteworthy that the utilization of eye-tracking in conjunction with head movement demonstrated improved performance. This outcome is sensible given that users retained the option to employ head movements for object selection, enhancing performance by approximately 11%. Importantly, this performance gain did not correspond to significant differences in terms of usability and task load, which is promising.

Both head and eye-tracking emerge as viable alternatives to Brain-Computer Interfaces (BCI) as showed in [65].
5.1.2 Selection and Manipulation Performance with Avatar

Interaction with the avatar instead of directly interacting with the object exhibited poorer performance, primarily due to the additional steps required before object manipulation became possible. No single input modality significantly outperformed the others concerning the sense of embodiment, task load or usability in the interaction task.

It is worth noting that eye tracking can be more demanding on the eyes, whereas head tracking can be more physically taxing. Future research could explore the effects of these input modalities for extended VR sessions.

5.1.3 Task Duration

One important aspect to consider is the brevity of the interaction task. With less than one minute of immersion at a time, users may not have had adequate time to acclimate to their virtual bodies, potentially mitigating the onset of negative symptoms, discomfort or sense of embodiment.

5.1.4 Level of Experience

With increasing experience in VR and video games, performance differences tend to diminish. Results for inexperienced individuals may improve with extended exposure to the simulation.

5.1.5 Comparison with Previous Studies

Our results deviate from expectations in some aspects. Despite one condition offering more control over the avatar, no significant difference in the sense of embodiment was found. This could be attributed to the limitations of the avatar or disparities between the virtual and real-world experiences that the Inverse Kinematics (Inverse Kinematics (IK)) presented to users. However, in the same way as [53] and [55], a strong sense of agency was still induced despite the users not moving their own body.

Additionally, we observed superior results with eye-tracking. This is somewhat contrary to what was found by [64] or [68] but it is crucial to consider that users could still use their heads in the eye-tracking condition, making it effectively an enhanced version of head tracking. Separating eye and head movement in VR is challenging, given that the viewpoint remains linked to head movements.
5.1.6 Conclusion

For the interaction task, considering the absence of significant differences and the slightly improved performance with eye-tracking, despite its susceptibility to bugs and errors, we recommend incorporating eye-tracking for users who cannot move their heads. However, it is not essential to ensure good performance and a correct sense of agency (around 5.0/7.0). Regarding task-load and usability, having control over the avatar only when selecting an object with the head seemed preferable, but the results lack significance, necessitating further research.

5.2 Research Question 2: Navigation

This section of the discussion centers on the outcomes associated with Research Question 2 1.2.1.1, which explores how various methods of navigation influence users’ sense of embodiment and usability in virtual reality. Research Question 2 aims to understand how different input modality and navigation techniques during a navigation task influences performance, usability, and the sense of embodiment.

5.2.1 Navigation Task Differences

In the navigation task, we observed significant differences in response to all questionnaires. These disparities extended to the time required to complete the task. It is essential to recognize that our experiment’s methodology introduced a potential bias. Participants encountered both tasks in the same order, which could have impacted their responses to questionnaires. For instance, participants may have provided more extreme responses towards the end of the experiment due to boredom or fatigue.

5.2.2 Navigation Methods

In the navigation task, participants appeared to invest considerable time in achieving precise pointing, which is not required in real-time navigation. This accounts for better completion times despite covering more distance, since participants commenced their movements toward the target directly.
5.2.3 Rotation in Navigation

Regarding the navigation task, the rotation squares were not customized for individual users. The introduction of a calibration step could potentially yield significantly improved results by ensuring that rotation controls are optimally positioned for each participant.

5.2.4 Level of experience

As users gain experience in VR and video games, performance differences tend to diminish, suggesting that results for inexperienced individuals may improve with prolonged exposure to the simulation.

5.2.5 Comparison with Previous Studies

Our findings suggest that steering through User Interfaces (User Interface (UI)) is feasible in VR, with the steering condition yielding better results than the trigger method, which is contrary to what was found by [76].

Similarly to [56], we did not find a significant impact of the locomotion technique on the sense of embodiment, despite our methods giving a different level of avatar control.

Additionally, head-tracking outperformed eye-tracking, consistent with prior research for navigation, such as the ones done in [20].

5.2.6 Conclusion

For the navigation task, the results were notably significant across various aspects, highlighting head-tracking with navigation steering as the optimal choice for performance. Furthermore, head-tracking is superior in terms of the sense of agency, usability, and task-load. The impact of navigation steering versus navigation triggering is not significant, indicating that both methods offer viable alternatives when it comes to the sense of embodiment, user experience or task-load.

5.3 Overall Implications and Broader Context

In this section, we will delve into the overarching implications of our findings from both Research Question 1 and Research Question 2. We will discuss
how different eye-tracking-based control methods of an avatar impact users’ sense of embodiment and the overall user experience in VR. Additionally, we will consider the broader implications of our research within the context of the current state of the field and potential societal impacts.

5.3.1 Sense of Embodiment in VR

Our study has shown that the integration of eye-tracking with avatar control does not significantly influence users’ sense of embodiment in virtual reality. It does mean that both eye-tracking and head-tracking can be viable alternatives, with a correct sense of embodiment. Moreover, while head-tracking remains a viable option, it is essential to recognize that the combination of eye-tracking and head movement enhances performance without significantly compromising usability or task load.

Overall, this work challenges the notion that physical body movement is essential for embodiment, opening the door to more inclusive VR experiences for individuals with limited mobility.

5.3.2 User Experience and Task Performance

In the context of Research Question 2, we observed significant differences in task performance and user experience based on navigation methods. Head-tracking with navigation steering emerged as the optimal choice for performance, sense of agency, usability, and task load.

The fact that navigation steering and navigation triggering did not significantly impact the sense of embodiment suggests that users can adapt to different control methods without compromising their sense of embodiment in the virtual world. This adaptability has practical implications for the design of VR applications, allowing developers to choose control methods based on performance and comfort without sacrificing the user’s sense of embodiment.

5.3.3 Broader Implications and Societal Impact

Beyond the immediate implications for VR design and interaction, our research has broader implications for the field of human-computer interaction (HCI). As VR technology becomes more accessible and integrated into various aspects of daily life, understanding how different control methods impact the user experience is crucial. This knowledge can inform the development of more inclusive and user-friendly VR applications, potentially benefiting
individuals with disabilities and expanding the reach of virtual reality in education, therapy, and entertainment.

In a societal context, our findings highlight the importance of considering individual differences in user preferences and abilities when designing VR experiences. Tailoring VR interfaces to accommodate diverse user needs can enhance the overall accessibility and usability of VR technology, ultimately making it more inclusive for a wide range of users.

5.4 Conclusion

Section 5.3 provided a comprehensive overview of how different eye-tracking-based control methods impact users’ sense of embodiment and the user experience in VR. We discussed the implications of our findings in the context of the current state of the field and the potential societal impact of our research. In the final section, we offer concise recommendations and insights for future research directions based on the outcomes of this study.

5.5 Limitations

In this section, we address limitations related to the methods and the experiment itself.

5.5.1 Avatar Realism and Implications

While our research primarily focused on comparing different levels of control methods in virtual reality, it’s important to acknowledge that avatar realism plays a significant role in users’ sense of embodiment, potentially impacting the outcomes of our study.

5.5.1.1 Lack of Facial Interaction

In our study, we did not incorporate facial interaction as a component of avatar realism. The absence of facial expressions and emotional responsiveness in avatars might have hindered participants’ abilities to fully feel embodied in the virtual characters they controlled.
5.5.1.2 Naturalness of Body Movements

The naturalness of body movements and gestures is another crucial aspect of avatar realism. While our setup included only a headset and an inverse kinematics plugin for avatar control, the limitations of this setup might have affected the fluidity and naturalness of avatar movements. Users typically expect avatars to move in a lifelike manner, and any deviations from this expectation can impact the sense of embodiment.

5.5.1.3 Avatar Customization

Although avatar customization and personalization are powerful tools for enhancing the sense of body ownership in VR, they also introduce complexities and potential trade-offs. The process of creating highly personalized, and realistic avatars can be time-consuming and technically challenging for both developers and users, as well as potentially leading to the uncanny valley. However, research has shown that the degree of personalization and realism in avatars positively correlates with the user’s sense of embodiment [79, 80] that we did not try to replicate as we used the avatars showed in 3.3.

In conclusion, while our research primarily focused on control methods in VR, it is crucial to recognize the potential influence of avatar realism on the sense of embodiment. Future studies in VR should consider incorporating aspects of avatar realism, such as facial interaction, natural body movements or personalized avatars, to gain a more comprehensive understanding of the factors contributing to user experiences in virtual environments.

5.5.2 Healthy Population

Our study exclusively involved healthy individuals, making it challenging to assess the potential impact on disabled individuals. Those accustomed to limited hand movement may experience heightened engagement and a higher Sense of Agency (SoA) when allowed to select and move their hands independently. Healthy individuals are familiar with controlling avatars or simulations with their entire bodies, whereas disabled individuals rely on alternative methods.
5.5.3 Participant Demographics

It’s important to note that the participants in our study were primarily engineering students from a university environment. This demographic may not be entirely reflective of the global population. Their familiarity with technical concepts and their comfort with technology might have influenced their ability to differentiate between the methods tested in our study.

Furthermore, the quality of the graphics in our virtual environment may not have met the expectations of participants with varying degrees of technical expertise. People less versed in technical matters might not discern subtle differences between the methods, especially if the visual quality was not on par with more advanced virtual reality experiences. This limitation should be considered when generalizing our findings to a broader population.

Future research should aim to include a more diverse participant pool, encompassing individuals from various educational backgrounds and technical proficiencies, to better understand the broader implications of our results.

5.5.4 No baseline

Omitting a condition where users employ their hands with controllers was a deliberate choice to avoid biasing the SoE in the following conditions. However, this decision complicates the assessment of our results in comparison to a baseline. Future research could include an experiment solely focused on individuals, using the baseline version to gather their perspectives. It would have been possible to add a third group doing only controllers’ method, but it was not feasible in the allowed timeframe.

5.5.5 Eye-Tracking Challenges

It is important to acknowledge occasional issues with eye-tracking, particularly in cases involving participants with large glasses or strong vision corrections. Calibration had to be repeated multiple times in such instances, affecting the initial user experience.

5.5.6 Level of Experience

Lastly, it is essential to recognize the limitation related to the level of experience categories, where some categories consisted of only two participants, potentially introducing bias into the results.
Moreover, for the VR experience impact, it seems that no effect was seen in the navigation task but considering that participants always experienced the selection & manipulation task first, once they started the navigation task, despite their little experience, they were not at level 0 anymore.

5.5.7 Conclusion

In conclusion, while our study is subject to certain limitations, we believe that the findings remain valuable for informing future research and design efforts. These limitations should be considered as opportunities for refinement and expansion in subsequent studies.
Chapter 6

Conclusions and Future work

This chapter presents conclusions and future work.

6.1 Conclusions

Returning to our initial objectives, our primary goal was to identify a viable method for controlling avatars in VR without relying on body movements and thoroughly evaluate its performance across multiple dimensions. We pursued this goal by proposing various methods for the two critical aspects of the VR experience: interaction (comprising selection and manipulation) and navigation. Regrettably, we did not obtain statistically significant results for the interaction component, making it challenging to provide specific recommendations in this regard. Nonetheless, our findings did yield actionable insights for the navigation task, where we recommend predominantly employing head-tracking, provided the user is capable of using it.

Regarding Research Questions 1 and 2, our study suggests that having more control over the avatar did not yield significant improvements in the SoE or usability for both interaction and navigation tasks. As for Research Question 3, it appears that head-tracking delivers superior results in every aspect for navigation, while eye-tracking only exhibits better performance in the interaction task.

In summary, while our study provides valuable insights, there are no definitive answers. Further research in this area is imperative.

In hindsight, we would allocate more effort to enhancing the realism of avatars and consider preparing a baseline version of the experiment with controllers for another group of participants. This would facilitate a more
comprehensive evaluation of the influence of IK, control methods, or input modalities on our results, allowing for broader applicability.

6.2 Future work

This section outlines potential directions for future research.

6.2.1 Baseline Comparison

As discussed in the previous section, it would be valuable to compare our methods against a baseline. Creating an additional group of participants who perform the experiment with controllers could provide critical insights into the impact of IK, methods, or input modalities on our results.

6.2.1.1 Exploring Face Interaction

Our study did not explore the possibilities of face interaction, which is a body part most quadriplegic individuals have control over. Investigating whether satisfactory performance and a stronger sense of embodiment can be achieved solely through face interactions or in combination with other modalities, such as eye-tracking or UI, holds potential for future research.

6.2.2 Adaptation to Existing Software

Adapting existing VR rehabilitation software to accommodate our system is another area worth exploring. This could involve assessing the compatibility of our methods with existing applications, making them widely accessible, and testing them in real rehabilitation scenarios.

6.3 Reflections

Despite the absence of significant results, our study contributes to the growing field of accessible VR, which is gaining increased attention. We take pride in our contribution, even though our results are preliminary. We view our work as a small step that may aid future research in developing more effective solutions. Furthermore, designing for a niche group often leads to the natural evolution of systems that benefit a wider audience, and even more so individuals with disabilities in our study.
References


[41] R. Fribourg, “Contribution to the study of factors influencing the sense of embodiment towards avatars in virtual reality,” These


Appendix A

Questionnaires

The questionnaires that were used.
A.1 Virtual Embodiment Questionnaire

Instructions
Please read each statement and check the relevant response to indicate how strongly you agree or disagree with each statement (1 through 7). There are no right or wrong answers. Answer spontaneously and intuitively.

<table>
<thead>
<tr>
<th>Acceptance/Body Ownership</th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Somewhat disagree</th>
<th>Neither agree nor disagree</th>
<th>Somewhat agree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC1 My body</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AC2 body/Parts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AC3 Human-like</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AC4 Belonging/Me</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cyber/Agency</th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Somewhat disagree</th>
<th>Neither agree nor disagree</th>
<th>Somewhat agree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO1 My movements</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO2 cookbook/Movements</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO3 control/Movements</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO4 own/Movements</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Change</th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Somewhat disagree</th>
<th>Neither agree nor disagree</th>
<th>Somewhat agree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH1 manual/shape</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CH2 achromatic/fit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CH3 achromatic/fit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CH4 achromatic/fit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure A.1: Virtual Embodiment Questionnaire
A.2 System Usability Scale

System Usability Scale


1. I think that I would like to use this system frequently

2. I found the system unnecessarily complex

3. I thought the system was easy to use

4. I think that I would need the support of a technical person to be able to use this system

5. I found the various functions in this system were well integrated

6. I thought there was too much inconsistency in this system

7. I would imagine that most people would learn to use this system very quickly

8. I found the system very cumbersome to use

9. I felt very confident using the system

10. I needed to learn a lot of things before I could get going with this system

Figure A.2: System Usability Scale Questionnaire
A.3 Nasa Task-Load-Index

NASA Task Load Index

Hart and Staveland’s NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

<table>
<thead>
<tr>
<th>Name</th>
<th>Task</th>
<th>Date</th>
</tr>
</thead>
</table>

Mental Demand  How mentally demanding was the task?  
Very Low  Very High

Physical Demand  How physically demanding was the task?  
Very Low  Very High

Temporal Demand  How hurried or rushed was the pace of the task?  
Very Low  Very High

Performance  How successful were you in accomplishing what you were asked to do?  
Perfect  Failure

Effort  How hard did you have to work to accomplish your level of performance?  
Very Low  Very High

Frustration  How insecure, discouraged, irritated, stressed, and annoyed were you?  
Very Low  Very High

Figure A.3: Nasa Task-Load-Index Questionnaire
A.4 Demographics Questionnaire

What is your previous experience with virtual reality? *

○ None
○ Very limited previous experience
○ Some previous experiences
○ Very familiar with virtual reality

What is your previous experience with video games? *

○ None
○ Very limited previous experience
○ Some previous experiences
○ Very familiar with video games

Figure A.4: Demographics Questionnaire used before the experiment
Appendix B

Data Graphs and Tables

In the following appendix, you will find all the results that are non-significant.

B.1 Performance

All the results related to the speed of a participant.

B.1.1 Interaction

In order to understand the following graphs, it is important to note that:

- Condition A: Object + Head
- Condition B: Avatar + Head
- Condition C: Object + Eye
- Condition D: Avatar + Eye
Figure B.1: Interaction Run Duration Depending on the coupled conditions
In order to understand the following graphs, it is important to note that:

- Condition A: Goal + Head
- Condition B: Avatar + Head
- Condition C: Goal + Eye
- Condition D: Avatar + Eye

Figure B.2: Interaction Total Duration Depending on the coupled conditions
Figure B.3: Navigation Run Duration depending on the order of the run
Figure B.4: Navigation Run Duration depending on the rotation direction
Figure B.5: Navigation Run Duration depending on the Coupled Conditions
B.2 Sense of Embodiment

All the results related to the Virtual Embodiment Questionnaire.
B.2.1 Interaction

Figure B.7: Acceptance depending on the Degree of Control and Input Modality
Figure B.8: Control depending on the Degree of Control and Input Modality

Figure B.9: Change depending on the Degree of Control and Input Modality
B.2.2 Navigation

Figure B.10: Acceptance depending on the Navigation Method and Input Modality
Figure B.11: Control depending on the Navigation Method and Input Modality

Figure B.12: Change depending on the Navigation Method and Input Modality
B.3 System Usability

All the results related to the System Usability Scale.

B.3.1 Interaction

System Usability Score depending on the degree of control and input modality

Figure B.13: System Usability depending on the Degree of Control and Input Modality
Figure B.14: System Usability depending on the Coupled Conditions
B.3.2 Navigation

System Usability Score depending on the navigation method and input modality

Figure B.15: System Usability depending on the Navigation Method and Input Modality
Figure B.16: System Usability depending on the Coupled Conditions

B.4 Nasa Task-Load-Index

All the results related to the Nasa Task-Load-Index.
B.4.1 Interaction

Figure B.17: Task Load Index depending on the Degree of Control and Input Modality
B.4.2 Navigation

Figure B.18: Task Load Index depending on the Navigation Method and Input Modality

B.5 Virtual Reality and Video game Experience Impact

All the results related to the Experience Questionnaire.

B.5.1 Interaction

Table B.1: Tukey - Interaction - VR Table

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>mean(A)</th>
<th>mean(B)</th>
<th>diff</th>
<th>se</th>
<th>T</th>
<th>p-tukey</th>
<th>hedges</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>5.0</td>
<td>4.0</td>
<td>1.1</td>
<td>2.0 × 10⁻¹</td>
<td>5.4</td>
<td>6.3 × 10⁻⁷</td>
<td>8.5 × 10⁻¹</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>5.0</td>
<td>3.7</td>
<td>1.4</td>
<td>2.0 × 10⁻¹</td>
<td>6.7</td>
<td>4.4 × 10⁻¹⁰</td>
<td>1.2</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>5.0</td>
<td>3.9</td>
<td>1.1</td>
<td>2.2 × 10⁻¹</td>
<td>5.1</td>
<td>3.6 × 10⁻⁶</td>
<td>7.7 × 10⁻¹</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>4.0</td>
<td>3.7</td>
<td>0.3 × 10⁻¹</td>
<td>1.5 × 10⁻¹</td>
<td>2.0</td>
<td>1.9 × 10⁻¹</td>
<td>2.8 × 10⁻¹</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>4.0</td>
<td>3.9</td>
<td>2.8 × 10⁻²</td>
<td>1.7 × 10⁻¹</td>
<td>1.7 × 10⁻¹</td>
<td>10 × 10⁻¹</td>
<td>2.3 × 10⁻²</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>3.7</td>
<td>3.9</td>
<td>-2.7 × 10⁻¹</td>
<td>1.7 × 10⁻¹</td>
<td>-1.6</td>
<td>4.0 × 10⁻¹</td>
<td>-2.5 × 10⁻¹</td>
</tr>
</tbody>
</table>
Table B.2: Tukey - Interaction - Video game Table

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>mean(A)</th>
<th>mean(B)</th>
<th>diff</th>
<th>se</th>
<th>T</th>
<th>p-tukey</th>
<th>hedges</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>4.8</td>
<td>4.2</td>
<td>6.0×10⁻¹</td>
<td>2.5×10⁻¹</td>
<td>2.4</td>
<td>8.5×10⁻²</td>
<td>4.7×10⁻¹</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>4.8</td>
<td>4.0</td>
<td>8.1×10⁻¹</td>
<td>2.4×10⁻¹</td>
<td>3.4</td>
<td>4.5×10⁻³</td>
<td>6.3×10⁻¹</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>4.8</td>
<td>3.7</td>
<td>1.1</td>
<td>2.4×10⁻¹</td>
<td>4.7</td>
<td>1.8×10⁻₅</td>
<td>9.4×10⁻¹</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>4.2</td>
<td>4.0</td>
<td>2.1×10⁻¹</td>
<td>1.7×10⁻¹</td>
<td>1.2</td>
<td>6.3×10⁻¹</td>
<td>1.7×10⁻¹</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>4.2</td>
<td>3.7</td>
<td>5.2×10⁻¹</td>
<td>1.7×10⁻¹</td>
<td>3.1</td>
<td>1.2×10⁻²</td>
<td>4.5×10⁻¹</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4.0</td>
<td>3.7</td>
<td>3.1×10⁻¹</td>
<td>1.5×10⁻¹</td>
<td>2.1</td>
<td>1.4×10⁻¹</td>
<td>2.7×10⁻¹</td>
</tr>
</tbody>
</table>

B.5.2 Navigation

Table B.3: Tukey - Navigation - VR Table

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>mean(A)</th>
<th>mean(B)</th>
<th>diff</th>
<th>se</th>
<th>T</th>
<th>p-tukey</th>
<th>hedges</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2.1×10¹</td>
<td>2.2×10¹</td>
<td>−8.7×10⁻¹</td>
<td>1.6</td>
<td>−5.4×10⁻¹</td>
<td>9.5×10⁻¹</td>
<td>−8.4×10⁻²</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>2.1×10¹</td>
<td>1.8×10¹</td>
<td>2.4</td>
<td>1.7</td>
<td>1.4</td>
<td>4.8×10⁻¹</td>
<td>3.9×10⁻¹</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>2.1×10¹</td>
<td>1.7×10¹</td>
<td>4.0</td>
<td>1.8</td>
<td>2.3</td>
<td>1.1×10⁻¹</td>
<td>6.9×10⁻¹</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>2.2×10¹</td>
<td>1.8×10¹</td>
<td>3.3</td>
<td>1.3</td>
<td>2.6</td>
<td>4.6×10⁻²</td>
<td>3.5×10⁻¹</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>2.2×10¹</td>
<td>1.7×10¹</td>
<td>4.8</td>
<td>1.3</td>
<td>3.6</td>
<td>2.1×10⁻¹</td>
<td>5.1×10⁻¹</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1.8×10¹</td>
<td>1.7×10¹</td>
<td>1.5</td>
<td>1.5</td>
<td>1.0</td>
<td>7.2×10⁻¹</td>
<td>2.8×10⁻¹</td>
</tr>
</tbody>
</table>

Table B.4: Tukey - Navigation - Video game Table

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>mean(A)</th>
<th>mean(B)</th>
<th>diff</th>
<th>se</th>
<th>T</th>
<th>p-tukey</th>
<th>hedges</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2.8×10¹</td>
<td>2.1×10¹</td>
<td>6.8</td>
<td>2.0</td>
<td>3.5</td>
<td>3.1×10⁻³</td>
<td>5.6×10⁻¹</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>2.8×10¹</td>
<td>1.9×10¹</td>
<td>8.9</td>
<td>1.9</td>
<td>4.7</td>
<td>2.2×10⁻⁵</td>
<td>1.0</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>2.8×10¹</td>
<td>1.8×10¹</td>
<td>9.7</td>
<td>1.8</td>
<td>5.3</td>
<td>1.0×10⁻⁶</td>
<td>1.2</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>2.1×10¹</td>
<td>1.9×10¹</td>
<td>2.0</td>
<td>1.4</td>
<td>1.5</td>
<td>4.5×10⁻¹</td>
<td>2.4×10⁻¹</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>2.1×10¹</td>
<td>1.8×10¹</td>
<td>2.9</td>
<td>1.3</td>
<td>2.2</td>
<td>1.2×10⁻¹</td>
<td>3.7×10⁻¹</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1.9×10¹</td>
<td>1.8×10¹</td>
<td>8.6×10⁻¹</td>
<td>1.2</td>
<td>7.2×10⁻¹</td>
<td>8.9×10⁻¹</td>
<td>1.5×10⁻¹</td>
</tr>
</tbody>
</table>